التحليل الإحصائي

التجــــاري

د. عزت قناوي

دكتوراه الفلسفة في الاقتصاد
والعلوم السياسية

د. نيرة سليمان

أستاذ الاقتصاد المساعد
بالمركز القومي للبحوث

دار العلم للنشر والتوزيع

٢٠٠٦

مقدمــــة

لا شك أن التغيرات الجذرية التي أحدثتها الثورة المعلوماتية والعصرنة التكنولوجية في جميع مناحي الحياة الاقتصادية والاجتماعية والعلمية أدت إلـى زيادة فاعلية الاعتماد على أسـاليب ووسـائل القيـاس والإحصـاء لمعالجـة المشكلات الناجمة عن افرازات هذه المعلوماتية بكل ما تحملـه مـن بيانـات ومعلومات لمسايرة ركب التقدم العلمي والحضاري . وفي ضوء ذلك كان مـن غير المفيد التعرض لمعالجة مشكلة معينة من خلال الحدس والتخمـين بـدون الاعتماد على الأسس والحقائق العلمية والتي يعتبر علم الإحصاء أحد دعائمهـا ومنارتها الفكرية ، نظراً لما يشتمل عليه من بيانات ومعلومات وأرقام مصنفة ومبوبة ومجدولة في إطار المشكلة المعنية بالبحث ثم تحليل وتفسير واختبـار هذه البيانات واستنباط المؤشرات والدلائل بغرض استجلاء الحقائق المرتبطـة بها وكيفية التنبؤ على أساسها في ضوء المعطيات والفروض التي يوفرها علم الإحصاء للوصول إلى اتخاذ القرار بناء على ركائز بحثية وعلمية سليمة .

وتدعيماً لهدف الكتاب فقد راعينا البساطة والوضوح وعدم التعـرض للمفاهيم الدقيقة بشأن النظريات الإحصائية التي تحتاج إلى خلفية كبيـرة مـن التحليل الرياضي ، لذلك فقد توخينا أن يغطي هذا الكتاب الموضوعات الأساسية ذات الأهمية للدارسين التجاريين من خلال تقديم النمـاذج والأمثلـة العمليـة المتعددة بما يساعد على سهولة فهم واستخدام هذه الموضـوعات فـي الحيـاة العملية .

والله من وراء القصد يلهمنا سواء السبيل ،،،

المؤلف

د/ عزت قناوي

فهرس الموضوعات

الفصل الأول
مفاهيم أساسية

- **تعريف علم الإحصاء :**

يعتبر علم الإحصاء أحد فروع علوم الرياضة التطبيقية حيث يختص بتجميع Collection وتقسيم واختصار Classification & Condensation وعرض Presentation البيانات الإحصائية وتحليلها Analysis of Data بحيث يمكن التعبير عن هذه النتائج في صورة جداول Tabular Form أو أشكال بيانية Graphic Form دون حذف مخل أو تغيير في بياناتها الأساسية مع استعمال بعض الاختبارات التي تؤدي إلى الوصول إلى استنتاجات وقرارات مناسبة.

وقد كان الإحصاء يستخدم بمعرفة الإحصائيين بهدف جمع البيانات الأساسية التي تعتمد عليها الدولة للتعرف على خصائص السكان مثل البيانات عن المواليد والوفيات وعدد السكان لأغراض التجنيد في الجيش وكذلك اهتمام الدولة بجمع بيانات عن الصادرات والواردات والدخول والإنفاق بغرض معرفة حجم التجارة الخارجية وفرض الضرائب ثم عرض هذه البيانات في صورة مختصرة كالجداول والأشكال البيانية واستخدام بعض المقاييس الرقمية البسيطة بدلاً من النظر إلى كل البيانات الخام في حجمها الأصلي لتفهم معنى هذه البيانات وهو ما كان يعرف بالإحصاء الوصفي.

ومع تطور الحياة تعددت المشاكل وكان لابد لعلم الإحصاء أن يتطور حتى يمكن استخدامه في حل هذه المشاكل ولذلك بدء استخدام الإحصاء بالتدريج في نواحي كثيرة بهدف الوصول إلى حقيقة الأمور واتجاهات الظـواهر المختلفـة سواء كانت اقتصادية أو اجتماعية أو علمية بحتة أو .... الخ. ولـم يثبـت أن انتشر استخدام علم الإحصاء في نواحي مختلفة حيث ظهرت فائدتـه كطريقـة سليمة من طرف البحث العلمي الدقيق ولم يلبث تطبيقه على النواحي التي تهتم بها الحكومات في تدبير سياستها وتصريف شئونها العامة بل تعداها إلى جميع الظواهر المختلفة، كذلك شئون الأفراد والهيئات الخاصة التي لا تمت الحكومة بصلة ما.

وينطوي علم الإحصاء على فرعين أساسيين هما :

(١) الإحصاء النظري : وهو عبارة عن مجموعة من القواعد والقـوانين والطرق الإحصائية التي تم التوصل إليها من خـلال جهـود علمـاء الرياضة الذين تخصصوا في علم الإحصاء وقاموا بتطوير نظرياتـه على أسس رياضية وعلمية بحتة.

(٢) الإحصاء التطبيقي : وهو ذلك الفرع الذي يختص بتطبيق النظريـات الإحصائية في كافة مجالات البحث العلمي. ويجب علـى المهتمـين بالمجالات التطبيقية أن يكونوا على دراية تامة بالنظريات الإحصـائية وطريقة تطبيقها وتفسير نتائجها. كما يجب الإلمام التام بأسس وقواعد وطبيعة المجالات التطبيقية موضع البحث وذلك حتى يتسنى لهم اختيار أنسب الأساليب الإحصائية اللازمة لجمع وتبويب وتحليـل البيانـات الإحصائية الخاصة بالمشكلة موضع البحث.

وعلى ذلك فإن الإحصاء يعتبر علم لأن له نظرياته وقوانينه التي تم التوصل إليها من خلال تطبيق طريقة البحث العلمي المنظم، كما أنه فن لأن نجاح تطبيقه يعتمد بالدرجة الأولى على خبرة الإحصائي بأسس المجال العلمي الذي تطبق فيه هذه النظريات الإحصائية.

- **أسلوب وخطوات البحث العلمي Scientific Method**

يعتمد الأسلوب العلمي في البحث على اتجاهين :

(١) **الاستنباط Deductive Method** : وتعتمد هذه الطريقة على الاستنتاج من العام إلى الخاص وهي تعتمد على التفسير التطبيقي وهو عبارة عن تحقيق أو تفسير ظاهرة من قانون أو ظاهرة عامة. ومثال ذلك تطبيق قوانين الوراثة على صفات النباتات، والحيوانات.

(٢) **الاستقراء Inductive Method** : وتعتمد هذه الطريقة على الاستنتاج من الخاص إلى العام ويتبع فيها ظاهرة التفسير الاستنتاجي وهي استنتاج ظاهرة عامة من مجموعة ظواهر خاصة كما حدث في استخلاص قوانين الوراثة بعد دراسة توزيع بعض الصفات في نسل النباتات.

وتعتبر كل من طريقتي الاستنباط والاستقراء هما أساس البحث العلمي الحديث. وتتلخص خطوات الأسلوب العلمي فيما يلي :

أولاً : تحديد المشكلة

لابد للباحث أن يقوم بتحديد المشكلة تحديداً دقيقاً وتفصيلياً، وتعتبر هذه الخطوة من أهم فترات البحث إذ يعتمد عليها في إمكانية الاستفادة من هذا البحث من عدمه. وعلى الباحث أن يعرف على قدر استطاعته كل ما يتعلق

بالمشكلة. من حيث نشوئها وأهميتها ونوع البيانات الضرورية وعن الطريق الذي ستستخدم فيه نتائج البحث.

ثانياً : تكاليف البحث

يجب أن يأخذ الباحث في الاعتبار تكاليف البحث والعوامل التي تؤثر في زيادة أو قلة هذه التكاليف، وعلى الشركة أو المؤسسة أن تقرر بعد معرفة تكاليف البحث والقيام به من عدمه.

ثالثاً : الفروض

يقوم الباحث بوضع فرض علمي يتصل بطبيعة المشاهدات في بحثه. وتختلف الفروض من حيث درجة بساطتها وتعقدها فقد يضع الباحث فرضاً بسيطاً وفي هذه الحالة لا يكون إلا مجرد تعميم أو تعبير سطحي للمشاهدات. أما الفروض الديناميكية أو المعقدة فتفترض وجود ترابط بين الحوادث وتأثيرها على بعضها الآخر. ولذلك فأن المشابهة أو التشابه من العناصر المهمة في إنشاء أو تكوين الفروض. ولذلك أيضاً فإن وضع الفروض يعتمد على براعة وخبرة الباحث نفسه.

رابعاً : جمع البيانات

ويتم ذلك عن طريق الملاحظة مع الاستعانة بالآلات الدقيقة والتجارب وكذا من المصادر الأخرى المختلفة.

خامساً : تجهيز البيانات وتصنيفها

يقوم الباحث بتجهيز البيانات المجموعة وتصنيفها حتى يمكن الاستفادة بها عن طريق استخدام الجداول والأشكال البيانية.

سادساً : تحليل البيانات واستخلاص النتائج

تستخدم الأساليب الإحصائية التحليل وتفسير البيانات واستخلاص النتائج والتي قد تكون من التعقيد بمكان إذا ما حاول الباحث وصفها باللغة العادية كما تساعده على تجنب أخطاء الاستنتاجات.

سابعاً : اختبار الفروض

حتى يستطيع الباحث تعميم نتائجه المستخلصة من بحثه بالنسبة للمجتمع يلزمه اختبار الفرض ونتائجه. ويستعين الباحث لأداء هذه المهمة بالأساليب الإحصائية المختلفة .

أقسام **علم الإحصاء Categories of Statistics :**

يمكن تقسيم علم الإحصاء إلى قسمين أساسين هما :

(١) الإحصاء الوصفي Descriptive Statistics : وهو ذلك القسم الذي يقتصر على جمع وعرض البيانات كما هي أو بعد استخراج بعض المؤشرات الإحصائية كالمتوسطات أو النسب التي تفسر تلك البيانات وذلك بغرض وصف الاتجاهات والخصائص المختلفة للظواهر موضع البحث.

(٢) الإحصاء الرياضي أو الاستدلال الإحصائي
Mathematical Statistics and Statistical Inference

وهو ذلك القسم الذي يقوم على تعميم نتائج الجزء على الكل وذلك من خلال دراسة عميقة من هذا المجتمع وتعميم النتائج المتحصل عليها على المجتمع كله وذلك من خلال تطبيق بعض أدوات التحليل مثل التوزيعات الاجتماعية واختبار الفروض الإحصائية.

أهداف علم الإحصاء :

نال علم الإحصاء في عصرنا الحديث اهتماماً متزايداً من الإحصائيين والهيئات الإحصائية في الدول المختلفة نظراً لإسهاماته العديدة في الكثير من التطبيقات الاقتصادية والاجتماعية والعلمية. نوجز بعضاً منها فيما يلي :

١) يهدف علم الإحصاء إلى توفير الإحصاءات السكانية Population Statistics وكذلك الحيوية اللازمة لرسم السياسات التخطيطية في شتى المجالات، منها على سبيل المثال : التعليم والصحة والرعاية الاجتماعية والأمن والعمالة. هذه الإحصاءات وما تتضمنه من بيانات إحصائية عن توزيع السكان حسب المناطق الجغرافية المختلفة وحسب العمر والنوع والحالة التعليمية، وكذلك الإحصاءات الحيوية من معدلات للمواليد والوفيات والخصوبة والإنجاب والزواج والطلاق هذا بالإضافة إلى توزيع قوة العمل (مشتغلين ومتعطلين) في المناطق المختلفة حسب المهن والأنشطة الاقتصادية، كل هذه الدراسات يجب توافرها أمام المخطط حتى يتسنى له معرفة الاحتياجات الخدمية لكل منطقة من المناطق، ومن هذا يستطيع وضع السياسات المناسبة لتحقيق الأهداف المرجوة في المجالات المختلفة . كذلك فإن البيانات الإحصائية الدقيقة والمعالجات الإحصائية السليمة لها تمكنان من تتبع سهو تنفيذ مراحل الخطط الموضوعة والحكم على مدى كفاءة كل مرحلة في تحقيق المستهدف لها من الخطط المرسومة.

٢) يسهم الإحصاء في تقديم الطرق والأساليب الإحصائية التي تستخدم في مراقبة وضبط الإنتاج Quality Control في المنشآت الصناعية ومن ثم الحكم على مدى مطابقة المنتج للمواصفات المقبولة من عدمه.

ولقد أصبح في معظم المصانع ما يسمى بوحدة مراقبة جودة المنتج حيث يوجد بها خرائط لمراقبة نسب المعيب أو متوسط الوحدات المعيبة، هذه الخرائط تحتوي على المستوى المقبول لهذه النسب أو المتوسط ومستويين أو حدين أعلى وأسفل المستوى المقبول يجب إلا تتجاوزه المستويات المشاهدة من العينة التي سحبت من إنتاج المصنع ومن ثم فالقرار بمدى مطابقة المنتج للمواصفات المقبولة دون توقف خطوط الإنتاج عن العمل واتخاذ القرار الإحصائي.

٣) أصبح تطبيق الأساليب الإحصائية واستخدام البيانات الإحصائية من أهم سمات البحوث العلمية في العلوم التجريبية كالزراعة والطب والصيدلة. فالتجربة أو التجارب التي تجري لتحقيق هدف ما أو اختبار فرض ما يجب أن تصمم بأسلوب يتناسب والهدف المرجو منها. وطرق تصميم مثل هذه التجارب هي مجال لفرع من فروع علم الإحصاء يسمى تصميم التجارب Design Of Experiments. هذا بالإضافة إلى أن تحليل نتائج هذه التجارب واتخاذ قرار برفض أو قبول الفروض التي يراد اختبارها، كل ذلك يتم عن طريق ما يسمى باختبارات الفروض الإحصائية Tests of Hypotheses والتي تعتبر ضمن موضوعات أخرى تكون فرع الاستدلال الإحصائي Statistical Inference وهو أحد مجالين رئيسيين لعلم الإحصاء أما المجال الأخر فيعرف بالإحصاء الوصفي Descriptive Statistics وهو المجال ذو الأهمية العظمى حيث يهتم الإحصاء الوصفي بجمع وتبويب وعرض جدولي وبياني للبيانات الإحصائية، كما يهتم بدراسة جانب من المقاييس الإحصائية التي تتعلق بتلخيص رقمي للظاهرة موضع الدراسة وأيضاً دراسة العلاقة بين ظاهرتين .

وجدير بالذكر أن عرضنا لما يمكن أن يقدمه الإحصاء في المجالات المختلفة قد أعطى لنا فكرة عن بعض فروع هذا العلم، لكن مع اتساع هذه المجالات وتنوعها أضيفت فروعاً جديدة مثل الإحصاء الرياضي Mathematical Statistics وهو فرع هام حيث بالربط بين الإحصاء والرياضيات ووضع الأسس الرياضية للإحصاء وتطويره.

وغني عن البيان أن كل ما تحتويه فروع المعرفة الإحصائية هو في إطار من النظرية الإحصائية Theory of Statistics والتي تحوي بين طياتها المادة العلمية للنظريات الرياضية والإحصائية وأسس الفكر الإحصائي. التي على أساسها يمكن تقديم أساليب التحليل الإحصائي المختلفة وطرق إمكانيات استخدامها في المجالات المختلفة.

علاقة الإحصاء بالعلوم الأخرى

أحتل علم الإحصاء مكانة بين العلوم الأخرى باعتباره الأساس في عمليات التخطيط وإجراء البحوث لدراسة الظواهر المختلفة لما يقدمه علم الإحصاء من بيانات موضوعية تساعد على معرفة حقيقة ومسببات أي ظاهرة، فلا يتصور إجراء بحث بدون استخدام الإحصاء وفيما يلي نوجز بعض استخدامات الإحصاء في المجالات المختلفة.

أ) الإحصاء والرياضيات :

نستطيع القول بأن علم الإحصاء قد أخذ الكثير من علوم ونظريات الرياضيات فقد نشأت خلال القرن التاسع عشر النظريات والطرق العلمية التي ما زالت إلى يومنا هذا تعتبر الركيزة الأساسية لعلم الإحصاء وكان ذلك نتيجة لتطور العلوم الرياضية تطوراً كبيراً خلال هذا القرن والقرن السابق عليه. ومع

ذلك الإحصاء علم حديث وإن كان قد اعتمد اعتماداً كلياً على تطـور العلـوم الرياضية في القرن التاسع عشر فإنه لم يظهر بصورة شـاملة وعامـة وبـدأ تدريسه في الجامعات والمعاهد العليا في أوائل القرن العشرين.

ب) الإحصاء والعلوم الاجتماعية :

تعتبر العلوم الاجتماعية هي المجال الخصب لعلم الإحصاء حيث تسـتخدم الإحصاء في مختلف البحوث الاجتماعية ومنها أبحاث التربية وعلـم الـنفس، والوسائل التعليمية، وأبحاث التعليم، والتعليم العالي، والسكان، وقوة العمل، كما يستخدم علم الإحصاء في حساب المعدلات والنسـب المسـتخدمة فـي كافـة المجالات الاجتماعية.

ج) الإحصاء والطب والزراعة والهندسة :

جميع الأبحاث الطبية تستخدم الإحصاء في تحليل النتائج التي تتوصـل إليهـا التجارب الطبية والزراعية والهندسية أيضاً والأمثلة على ذلك كثيـرة فحـين اختيار نوع معين من أنواع السماد أو كفاءة آلة من الآلات لابد مـن اسـتخدام الإحصاء للحكم على نتائج التجارب من بعد استخدامه أولاً في تصـميم هـذه التجارب كما يستخدم الإحصاء بشكل كبير وواسع في مجـال مراقبـة جـودة الإنتاج.

د) الإحصاء والعلوم الأمنية :

يعتبر الأمن حاجة فطرية لصيقة بالوجود الإنساني ويرتبط ارتباطـاً وثيقـاً بحاجاته وتطلعاته وتتأكد أهميته عندما تتعرض حاجة الإنسـان الأمنيـة لأي صعوبات على أن المفهوم العصري للأمن اتسع كثيراً بحيث يشمل نشاط الدولة في جميع المجالات، ولم يعد ذلك المفهوم الذي يقتصر على حماية المواطن من شرور الجريمة ومكافحتها قبل وبعد وقوعها بل أصبح يعنـي تـأمين مسـيرة

المجتمع نحو التحسين المستمر لرفاهية أفراده، وبهذه الدلالات فإن الأمن العام يمتزج امتزاجاً لا يمكن فصله عن أنشطة المجتمع. بكافة صورها، فهو الذي يمنح الإحساس بالراحة والاستقرار الذي يشعر به الأفراد تجاه أنفسهم وأموالهم وأعراضهم ومستقبلهم، وكذلك استقلالهم وتحررهم من كافة صور الاستغلال وبذلك فإنه يشكل مزيجاً من الأمن الجنائي والنفسي والاجتماعي والسياسية والحربي والاقتصادي وأي خلل في هذه المكونات يؤثر تأثيراً مباشراً على باقي المكونات وبالتالي على أمن المجتمع بأكمله.

إن هذه المكونات للأمن العام لا تختلف من مجتمع لآخر، ولكن الوزن النسبي لها يختلف من دولة إلى أخرى في الزمان الواحد كما يختلف داخل الدولة الواحدة في الأزمنة المختلفة تبعاً لمراحل التنمية التي يمر بها، ولذا فقد أصبح حفظ الأمن بجميع أشكاله ضرورياً ومهماً، لأن التنمية في جميع المجتمعات يجب أن ترتكز على قواعد ثابتة يتفاعل بها التخطيط مع التنفيذ وتتوالى الرقابة وتستمر على مواقع العمل لإنجاز المشروعات التنموية، وتحقيق ذلك مرهون باستتباب الأمن واستقرار الأوضاع السياسية والاجتماعية والاقتصادية، والشرطة بأجهزتها المتنوعة هي الهيئة المنوطة بها توفير الأمن وإقرار النظام وحماية المجتمع من الجرائم التي تطورت وتعقدت مع التقدم الحضاري والتكنولوجي الأمر الذي يجعل في مقدمة الأولويات توجيه أقصى الرعاية والاهتمام برجال الشرطة عن طريق الاستمرار في تنمية قدرات ومهارات العاملين بها وتعميق معارفهم وخبراتهم بالإضافة إلى تزويدهم بأحدث الأجهزة ليتسنى لهم إنجاز المتطلبات الوظيفية بكفاءة وفاعلية، وليس باستطاعة أي جهاز للأمن أن يتطلع إلى تحقيق أهدافه بكفاءة عالية إلا بالأخذ بأساليب

التخطيط العلمي وأسس التنظيم المتطورة والقيادة الرشيدة واتخاذ القرارات بأساليب التخطيط العلمي وأسس التنظيم المتطورة والقيادة الرشيدة واتخاذ القرارات بالأسلوب العلمي وأساليب الرقابة الإيجابية.

تحقيقاً لهذه المفاهيم احتل علم الإحصاء الجنائي مكانته باعتباره الأساس في عمليات التخطيط ودراسة الظواهر المختلفة لما يقدمه من بيانات موضوعية تساعد على معرفة حقيقة ومسببات أية ظاهرة كما يمهد السبيل لتحليل وتقويم الحالة الأمنية وتفسيرها وإعطاء الحلول المناسبة لها معتمداً على لغة الأرقام التي أصبحت الأساس السليم لأي قرار في جميع الحالات وعلى مختلف الأصعدة.

مسئوليات الإحصائي

إن تقدم العلوم وما حققته التكنولوجيا الحديثة في كل فرع من فروع المعرفة بالقطع استفادت منه البشرية .

ولكن عندما نعتبر العلاقة بين الناس والإحصاء فإنه من الضروري اعتبار هذه العلاقة في اتجاهين أي هناك اتجاهين للتأثير. فعندما تستخدم النتائج والخلاصة الإحصائية في صنع القرارات التي تؤثر في حياة الناس فإن المعلومات والبيانات التي بنيت عليها هذه النتائج والتي توصلنا إلى القرارات عن طريقها أعطيت بواسطة الناس ويشارك الإحصائي في كلا الاتجاهين في جمع البيانات وفي استخلاص النتائج المبينة على هذه البيانات ليقدمها لنا كمنتج نهائي.

وبهذا فهو عليه نوعين من المسئولية :

- المسئولية الأولى : لمستخدم منتجه النهائي.

- المسئولية الثانية : للناس حيث قدموا له المادة الخام.

والإحصائي كمنتج للمقاييس أصبح حرصه يزداد تجاه هذه المسئولية وذلك لأن الإحصاء بصفة عامة والمسح بالعينة بصفة خاصة أصبح أداة هامة من أدوات صنع القرارات. فالقرارات التي تصنع على مستوى الحكومة والقطاع العام تقرر مثلاً ما إذا كان ينبغي الاستمرار في برنامج حكومي معين أو إنهاؤه وهل من الأفضل زيادة الدعم لمنتج أو نشاط معين أو تخفيضه أو إلغاؤه هذه القرارات يتم اتخاذها باستخدام المعلومات الإحصائية التي يتم جمعها باستخدام أسلوب العينات ومن هنا يكون ضرورياً على الإحصائي التزام الأسلوب العلمي السليم للوصول إلى نتائج تكون أساس لقرار حكيم.

وفي القطاع الخاص : فإن صنع القرار بالاعتماد على مسح لشواهد واتجاهات أصبح منتشراً للغاية. فقرار تقديم إنتاج جديد غالباً ما يبنى على قابلية المستهلك لهذا المنتج .

وفي مجال الصحة والتعليم : فإن صناعة القرار إحصائياً تستخدم بشكل مكثف كالعلاقة بين التدخين والسرطان وكالعلاقة بين درجة الذكاء ونتائج الامتحانات، وكلنا معاينين لهذه الأمثلة التي توضح استخدامات الإحصاء.

المصطلحات الإحصائية :

قبل أن نستطرد في دراستنا لمبادئ علم الإحصاء يجب التعرف على بعض المصطلحات الهامة التي سوف يرد ذكرها مراراً أثناء التقدم في الدراسة بغية تفهمها والتعرف على ماهيتها ومواصفاتها، وفيما يلي وصف مختصر لبعض هذه المصطلحات.

(أ) مجتمع العينة Population :

والمقصود به مجموعة من الأفراد أو الأشياء التي تشترك في صفة أو أكثر تشير إلى انتمائهم إلى مجموعة واحدة مثل طلاب الجامعات يشكلون مجتمعاً على اختلاف جنسهم وأطوالهم وأعمارهم وربما أجناسهم لأنهم يشتركون في صفة أو أكثر كالمستوى الثقافي قبل دخول الجامعة والهدف من دخولها، وقياساً على ذلك يمكن القول أن عمال صناعة ما يشكلون مجتمعاً، وأن عمال الزراعة يشكلون مجتمعاً أخر، وأن طلاب الابتدائي يشكلون مجتمعاً ثالثا وهكذا . وكلما تكرر القياس وازداد عدده كلما مثل مجتمعه أصدق تمثيل.

والمجتمع إما أن يكون محدود العدد مثل عدد الحاصلين على درجة الدكتوراه في مصر . وقد يكون غير محدود العدد كعدد النباتات الفطرية في العالم كله.

مثل هذا المجتمع يطلق عليه المجتمع النظري، وفي دراستنا الإحصائية لأي مجتمع يكون اهتمامنا بالمواصفات العامة التي تشكل وتكون من هذا المجتمع مجتمعاً بصرف النظر عن الاختلافات الفردية بين وحداته.

ب) العينة Sample :

غالباً ما يلجأ الباحث إلى اختيار عدد صغير من أفراد المجتمع الذي يدرسه بطريقة معينة بحيث تمثل هذا الاختيار أفراد كله بأكبر قدر من الثقة ذلك أنه إذا أراد دراسة أفراد المجتمع كله فربما قصرت مجهوداته ووقته وأمواله عن إمكان استكمال هذه الدراسة، لذا اعتبر الباحثون العينة أكفء وسيلة لدراسة المجتمع كله إذا ما توفرت لها صفات اختيارها ممثلة للمجتمع.

ولاشك في أن النتيجة التي يتوصل إليها الباحث من دراسة المجتمع كله أدق بكثير مما لو قام الباحث بدراسة عينة من هذا المجتمع، إلا أن الباحثين يضطرون إلى عدم دراسة المجتمع كله بل يكتفون بدراسة جزء منه كعينة للأسباب التالية :

١. توفير الجهد والوقت والمال.
٢. ربما كانت العينة هي الوسيلة الوحيدة الممكنة لدراسة المجتمع خاصة إذا كان غير محدود.

لذلك يجب إلا يخفى عن الأذهان أنه كلما زاد حجم العينة كلما كانت أكثر صدقاً في تمثيلها للمجتمع، وتختلف طريقة أخذ العينة من المجتمع باختلاف نوع البحث والغرض منه ونوع المجتمع وطبيعة أفراده ومدى الدقة المطلوب إتباعها، ونوع النتائج المطلوب الحصول عليها.

ويمكن تقسيم العينات الشائع استخدامها في التجارب الإحصائية إلى :

١. **العينة العشوائية البسيطة** Simple Random Sample

وهي التي يؤخذ أفرادها بطريقة تعطي فرصاً متكافئة لاختيار كل فرد فيها فمثلاً إذا أريد اختيار خمسين طالب من الخمسمائة التي تمثل هذا المجتمع، وتم إعطاء كل طالب رقماً وسجل كل رقم في بطاقة صغيرة متماثلة الحجم والسمك بحيث يتم اختيار الخمسين طالب بطريق المصادفة التامة أو باستخدام جداول

الأعداد العشوائية فإن هذه العينة تصبح ممثلة لمجتمع الخمسمائة طالب ويطلق على هذه الطريقة طريقة أخذ العينة العشوائية البسيطة.

٢. **العينة المتحيزة (غير العشوائية)** Non-Random Sample:

وهي التي تؤخذ بطريقة غير عشوائية أو متحيزة، وقد يرجع هذا التحيز إلى عدم خبرة الباحث بأخذ العينة العشوائية، وبالتالي تصبح نتيجة البحث غير صادقة للواقع، وقد يكون التحيز اضطرارياً كما يحدث عند اختيار الباحث لعينة صغيرة جداً من مجتمع كبير جداً، أو دراسة الحالة الاجتماعية في أحد الواحات في صحارى العالم، واعتماد البحث على هذه العينة يجعل النتيجة غير ممثلة للواقع، وكلما زاد عدد أفراد هذه العينة غير المتحيزة كلما قربت هذه النتائج من واقع الحقيقة في المجتمع المأخوذة منه، وعموماً فإن إجراء البحث بعينات متحيزة يؤدي في أغلب الأحيان إلى الحصول على نتائج غير ممثلة ولا ينصح كثيراً بإتباعها.

٣. **العينة الطبقية العشوائية** Stratified Random Sample :

وهي عبارة عن أخذ عينة عشوائية بسيطة من كل طبقة من طبقات المجتمع الذي يتصف بالطبقية، ومعنى الطبقية في المجتمع انقسامه إلى أجزاء تتصف أفراد كل منها ببعض الصفات المميزة عن صفات أجزاء القسم الآخر، ويكون مجموع العينات البسيطة العشوائية المأخوذة من كل طبقات المجتمع ممثلة بنسبتها في هذا المجتمع هي العينة الطبقية العشوائية الممثلة لكل المجتمع تمثيلاً دقيقاً.

فللحصول على عينة ممثلة للحيازات الزراعية الموجودة في منطقة ما أو في الجمهورية مثلاً يجب استخدام العينة الطبقية العشوائية حتى تصبح العينة المأخوذة ممثلة تمثيلاً صادقاً للواقع، وبها تقسم الحيازات إلى أقسام (أقل من ٥ أفدنة) (من ٥ إلى اقل من ١٠ أفدنة)، (من ١٠ إلى اقل من ١٥ فداناً) وهكذا ثم يؤخذ من كل مجموعة أو طبقة عينة عشوائية حجمها ممثلاً لنسبة حجم هذه المجموعة من المجتمع كله، ويكون مجموع هذه العينات العشوائية البسيطة هو حجم العينة الطبقية العشوائية المطلوبة.

المتغيرات Variables :

المتغير عبارة عن مجموعة من الصفات أو القيم يطلق عليها مجال وقد تمثل قيمة أو صفة واحدة فيطلق عليها ثابت، وقد يكون المتغير أو المجال أو الصفة المتغيرة وصفية كالذكاء، والاستيعاب واللون وغيرها وهو مالا يمكن قياسه، كما يكون المتغير أو المجال أو الصفة المتغيرة كمية وهي التي يمكن قياسها بمقاييس مختلفة كالمساحة والأطوال والأوراق ... الخ، والأخيرة تنقسم إلى قسمين : الأول : متغيرات وصفية كمية متصلة (مستمرة)، والثاني : متغيرات متقطعة أو غير مستمرة.

القيم الإحصائية Variates :

وهي عبارة عن التعديلات القيمية التي تحدث للظاهرة أو الصفة أو المجال أو التغير من فرد لآخر في مجتمع البحث ويطلق على كل قيمة في هذا المتغير رمز (س) يضاف إليه رقم عددي لتخصيص الترتيب مثل ($س_١$)، ($س_٢$)، ($س_٣$) ويعبر عن مجموع قيم المتغير بالرمز ($س_ن$) أما مجموع القيم الإحصائية في

مجال متغير أو صفة معينة فيطلق عليه رمز (مجـ س) وغالباً ما يظهر في العرض البياني على المحور السيني أو الأفقي .

عناصر الأسلوب الإحصائي :

يمكن تلخيص وظيفة علم الإحصاء أو ما يعبر عنه بعناصر الطريقة الإحصائية والأسلوب الإحصائي فيما يلي :

- جمع وترتيب البيانات.
- تبويب البيانات.
- عرض البيانات في شكل :
  1. سرد
  2. جداول
  3. رسوم إيضاحية أو بيانية
- تحليل البيانات واختبار معنويتها.
- تفسير النتائج بغية الوصول إلى ما يترتب عليها.

وأما الصفات المميزة للطريقة الإحصائية أو الأسلوب الإحصائي فهي :

- هي الوسيلة الوحيدة للتعامل مع الكميات الكبيرة للمعلومات الرقمية.
- هي طريقة تستخدم فقط في حالة ما إذا كانت المعلومات ممكن اختصارها إلى حالة كمية.
- أسلوبها واحد فالمستخدم في العلوم الاجتماعية، والتربية والاجتماع وعلم النفس، هو نفسه المستخدم في علم االأحياء والكيمياء، والفلك على الرغم من اختلاف طبيعة هذه المواد.

الفصل الثاني

المتغيرات الإحصائية

- **المتغير (Variable) :**

مقدار يأخذ العديد من القيم فالمقدار الذي يمكن أن تكون له أكثر من قيمة واحدة يمكن اعتباره متغير، فكلمة متغير تأتي من التغير (والتغير معناه عدم الثبات وعكس كلمة متغير ثابت Constant والثابت هو المقدار الذي يأخذ قيمة واحدة فقط لا تتغير أما المتغير فكما قلنا مقدار يأخذ العديد من القيم) فالطول مثلاً متغير يتغير من شخص لأخر فهذا طوله ١٦٠ سم وذاك طوله ١٦٥ سم وثالث طوله ١٧٠ سم وهكذا، والدخل أيضاً متغير فهذه الأسرة دخلها ١٠٠ جنيه شهرياً وتلك دخلها ٢٠٠ جنيه وثالثة دخلها ٥٠٠ جنيه وعادة ما نرمز للمتغير بالرمز س = ١٠٠ أو ٢٠٠ أو ٥٠٠ وهكذا حسب دخل الأسرة المعنية بالقياس في هذه الحالة.

أما الثابت فمقدار لا يتغير مثل ط = $\frac{٢٩}{٧}$ (النسبة التقريبية) ومثل درجة غليان الماء ١٠٠ درجة وأيضاً عدد ساعات اليوم ٢٤ ساعة وهكذا.

فللمتغير خاصية يمكن أن تكون لها أكثر من قيمة واحدة محددة من بين عديد من القيم أو الصفات، وكلمة متغير لا تعني بالضرورة أرقام بل تعني أحياناً صفات وأحسن مثال على ذلك المتغيرات في العلوم الاجتماعية، فالجنس

مثلاً متغير ق يأخذ القيمة ذكر أو القيمة أنثى والمستوى الاجتماعي متغير قد يأخذ القيمة غني أو متوسط أو فقير ... الخ.

والحالة الزواجية أيضاً متغير قد يتم توصيفه بأخذ القيمة متزوج، أعزب، مطلق، أرمل ... اخ.

والديانة أيضاً متغير يحمل الصفات مسلم، مسيحي، يهودي ... الخ، فعلى ذلك يتبين لنا أن المتغير لا يعني بالضرورة أن يأخذ صورة رقمية أي أن يكون في شكل عدد من الأعداد أو كسر من الكسور ولكنه يمكنه أن يأخذ قيماً وصفية فعلى عكس المتغيرات الرقمية كالطول الذي يأخذ قيماً تقاس بالسنتيمتر أو البوصة أو الوزن الذي يأخذ قيماً تقاس بالكيلو جراماً أو الباوند والدخل الذي يأخذ قيماً تقاس بالجنيه أو الدولار أو الريال نجد المتغيرات الوصفية التي تأخذ قيماً وصفية كما أوضحنا في حالات الجنس والمستوى الاجتماعي والحالة الزواجية والديانة ... الخ.

المتغير المستقل والمتغير التابع

في بعض الدراسات تستخدم الطرق الإحصائية في تحليل البيانات عن متغير واحد فقط والتحليل هنا يسمى أحادي المتغير (Unvaried) حيث يشتمل التحليل في هذه الحالة على متغير واحد فقط. قد يكون الهدف تحليل بيانات الأجور لمجموعة من العمال في مصنع معين أو تحليل بيانات الوزن لمجموعة من فئران التجارب في معمل أو تحليل بيانات المحصول في بعض القطع من أراضي التجارب ... الخ.

وفي دراسات أخرى تستخدم الطرق الإحصائية في تحليل البيانات عن متغيرين لدراسة العلاقة بينهم والتحليل هنا يسمى ثنائي المتغير (Bivariate)

حيث يشتمل التحليل هنا على متغيرين اثنين مثال ذلك في دراسة العلاقة بـين الكمية المعروضة من سلعة ما وسعرها أو الكمية المطلوبة من سلعة أخرى أو حتى نفس السلعة وسعرها. أيضاً عند دراسة العلاقة بين عمر طفل ومسـتوى ذكاؤه أو بين عمر الطفل ووزنه أو طوله أو محيط رأسه ... الخ. فهنا التحليل يشمل المتغيرين وتظهر هنا مشكلة تحديد المتغير المستقل (Independent V) والمتغير التابع (Dependant V) فحين كنا نعرض الدراسة لمتغير واحد فقط لم تكن لدينا هذه المشكلة لأن التحليل يشتمل على متغير واحد فقط ولكن حـين اشتمل التحليل على متغيرين فإنه لابد من تحديد أيهما أسبق على الأخر وبعبارة أخرى أيهما المتغير المستقل وأيهما المتغير التابع لأن ذلك سيبني عليه أسـاس التحليل.

والمتغير المستقل هو المتغير الذي تتغير القيم التي يأخذها بعيداً عن المتغير الأخر ودون أن يكون للمتغير الثاني أي تأثير على القيم التـي يأخـذها هـذا المتغير. أما المتغير التابع فذلك الذي تتوقف القيمة التي يأخذها هـذا المتغيـر على القيمة التي يأخذها المتغير الأول (المستقل) أي أن تغير قيمة هذا المتغيـر يكون نتيجة لتغير قيمة المتغير الأول فالعلاقة هنا سبب ونتيجة وعادة ما يرمز للمتغير المستقل بالرمز (س) وبالتابع بالرمز (ص) وأوضح مثال لذلك حـين دراسة العلاقة بين الكمية المعروضة لسلعة ما وسعرها فإن الكمية المعروضـة هنا هي المتغير المستقل والسعر هو المتغير التابع، فمن المعـروف أنـه فـي ظروف الحرية الاقتصادية عندما تزداد الكمية المعروضة من سلعة مـا فـإن سعرها ينخفض، فانخفاض السعر هنا نتيجة لزيادة المعروض من هذه السـلعة أي أنه عندما تغيرت الكمية المعروضة من السلعة من مقدار معين إلى مقـدار أكبر فإن السعر تغير إلى سعر أقل.

وبالمثل أيضاً حين دراسة العلاقة بين الكمية المطلوبة من سلعة ما وسعرها فإن الكمية المطلوبة من السلعة هي المتغير المستقل والسعر هو المتغير التابع لأنه أيضاً من المعروف اقتصادياً أنه عندما يزداد الطلب على سلعة ما فإن سعرها يرتفع وبالتالي فإن ارتفاع الطلب على سلعة ما يؤدي إلى ارتفاع سعرها وانخفاض الطلب على ذات السلعة يؤدي إلى انخفاض سعرها فارتفاع السعر هنا وانخفاضه جاء نتيجة لزيادة أو نقص الكمية المطلوبة.

وفي دراسات ثالثة تستخدم الطرق الإحصائية في تحليل البيانات عن أكثر من متغير لبيان العلاقة بينهما والتحليل هنا يسمى متعدد المتغيرات (Multivariate Analysis) حيث يشتمل التحليل في هذه الحالة على عديد من المتغيرات وعادة ما يكون أحدهم متغير تابع والباقي متغيرات مستقلة ويحدث ذلك حين يكون التغير الذي يحدث في متغير ما يتوقف على مجموعة من العوامل أدت في مجموعها إلى هذا التغير. هناك مثال لذلك كمية المحصول الناتج في زراعة معينة (المتغير التابع) هي نتيجة لمجموعة من العوامل مثل نوع التربة ونوع السماد وكمية وفترة الزراعة وعدد مرات الري وكمية المياه في كل مرة ومقدار الرعاية والعناية المبذولة خلال فترة الزراعة ... الخ وهذه كلها متغيرات مستقلة.

مثال أخر قد يعزي تفوق طالب معين في مادة ما لعديد من العوامل مثل عدد ساعات المذاكرة ودرجة ذكاؤه ومستوى الامتحان ووسيلة التدريس ومستواه الاجتماعي ودرجة تعليم الوالدين ... الخ. وبالتالي فإن الدرجة التي حصل عليها الطالب في هذا الامتحان لتلك المادة هي المتغير التابع وجميع العوامل الأخرى تعتبر متغيرات مستقلة. وعادة ما يرمز للمتغير التابع بالرمز (ص) والمتغيرات المستقلة بالرموز $س_١$، $س_٢$، $س_٣$، $س_٤$ ... الخ. ولا يعني ذلك أنه

في جميع الدراسات متعددة المتغيرات يجب أن يكون هناك متغير تابع واحد فقط وعديد من المتغيرات المستقلة بل قد ينسحب التحليل إلى مجموعة من المتغيرات المستقلة ومجموعة أخرى من المتغيرات التابعة ويرمز في هذه الحالة للمتغيرات المستقلة بالرموز س١، س٢، س٣، س٤ ... الخ. والمتغيرات التابعة بالرموز ص١، ص٢، ص٣، ص٤ ... الخ. وهذه من أصعب أنواع التحليل الإحصائي لأنها تتطلب خلفية رياضية قوية للغاية بالإضافة إلى أساليب خاصة في التحليل.

مستويات القياس

بعد التعرف على معنى كلمة متغير وعلى خصائص المتغير وطبيعته من حيث إمكانية كونه وصفياً أو رقمياً أو بمعنى أخر كيفياً أو كمياً (Quantitative or Qualitative).

فإنه يجدر التفرقة بين فرعين أساسيين من المتغيرات الكمية وهما المتغيرات المنفصلة (Discrete Variables) والمتغيرات المتصلة (Continuous Variables) فالمتغيرات المنفصلة هي المتغيرات التي تأخذ قيماً صحيحة فقط ولا تأخذ أية قيمة كسرية.

مثال ذلك عدد التلاميذ فقد يكون عدد التلاميذ في الفصول كالتالي ١٥، ٢٢، ٢٧، ٢٠، ٣٢ ... الخ وفي هذه الحالة لا يمكن أن يكون عدد التلاميذ في فصل معين ٢٥,٥ أو ٢٣,٧ أو ٣٥,٩١ ... الخ أو كمية الإنتاج المقام في مصنع للملابس الجاهزة فقد يكون الإنتاج اليومي ٥٠ قميص أو ١٢١ قميص أو ٢٤٧ قميص ... الخ، فأيضاً لا يمكن أن يكون الإنتاج ٢٣,٠٧ قميص أو ٤١٥,٣٨ قميص ... وهكذا.

أما المتغير المتصل فهذا الذي يمكن أن يأخذ أي قيمة سواء كانت صحيحة أو كسرية.

مثال ذلك الوزن أو الطول أو الدخل ... فيمكن أن يكون وزن قفص البرتقال مثلاً ٢٠ كيلو أو ٢١,٥ كيلو أو ٢٢,٢٥ كيلو ... الخ، ويمكن أن يكون طول الشخص ١٦٠ سم أو ١٦١,٣ سم أو ١٦٥,٦ سم ... الخ، والدخل من الممكن أن يكون ٣٥٢ جنيه أو ٣٥٢,٧ جنيه أو ٣٥٢,٦١ جنيه أو حتى ٣٥٢,٦١٨ جنيه ... وهكذا.

وهناك نقطة أساسية وهامة في التحليل الإحصائي للمتغيرات وهي أن أسلوب التحليل المناسب لمجموعة من المتغيرات يتوقف على طبيعة هذه المتغيرات والطرق التي قيست بها أو بمعنى آخر مستويات القياس (Levels of Measurement) والتي يمكن تعريفها من خلال أربعة مستويات أساسية لقياس المتغيرات نتعرض لهم واحداً بعد الآخر.

١) **المقياس الأسمي** (Nominal Scale) :

هناك العديد من المتغيرات التي تقاس بالمقياس الأسمي وهذا المقياس يستخدم للبيانات الكيفية (Qualitative Sata) فهو مجموعة من الأصناف التي تتباين كيفياً وليس كمياً أي تتغير في الصفة وليس في القيمة.

مثال ذلك حالة الزواجية لشخص ما قد يكون متزوج أو أعزب أو مطلق أو أرمل ... الخ. وأيضاً التمييز بين مجموعة من الأشخاص بناءاً على محل الميلاد : القاهرة، الإسكندرية، الشرقية، البحيرة، المنوفية ... الخ.، أيضاً التفرقة بين عدد من الأقمشة تبعاً لألوانها : أحمر، أخضر، أصفر، برتقالي ... الخ. وهذا المقياس هو أدنى مستويات القياس ولا يعدو أن يكون مجرد تمييز

للتفرقة بين مجموعات مختلفة بناء على صفة من صفات هذه المجموعات ولا يعني هذا المقياس مطلقاً أفضلية أو ترتيب تصاعدي أو تنازلي.

٢) المقياس الترتيبي (Ordinal Scale) :

هو قيم ترتيبية غير رقمية وهذا المقياس هو أعلى درجة من المقياس السابق فحيث أنه أيضاً يعتبر تميز للتفرقة بين الأشياء لكنه يعني بالإضافة إلى التمييز أفضلية ذات ترتيب تصاعدي أو تنازلي، ولكن ليس هناك مجال لتقدير المسافة بين كل فئة والفئة التي تليها ولا يعني تساوي المسافات بين فئة والفئة التي تليها تساوي الظاهرة. أي أن الفروق ليست ذات معنى، مثال ذلك قياس اتجاه الأشخاص نحو قضية معينة أو قياس رأيهم تجاه مشكلة محددة فقد يكون الشخص :

موافق جداً   موافق   غير محدد   غير موافق   غير موافق تماماً

أو عند ترتيب مجموعة من التلاميذ فتقول الأول والثاني والثالث ... الخ، فليس بالضرورة أن يكون الفرق بين الأول والثاني مساو للفرق بين الخامس والسادس أو بين الحادي والعشرون والثاني والعشرون ولكن نعلم أن الأول أفضل من الرابع والعاشر أفضل من الخامس عشر ... الخ.

ويستخدم هذا المقياس في عديد من البحوث الاجتماعية حيث يكون الهدف قياس ظواهر معينة عن طريق إبداء الرأي من قبل المبحوث في صورة :

تأييد تام   تأييد   غير محدد   رافض   رافض بشدة

فرأي المبحوث متغير يأخذ أحد الصفات الخمسة السابقة وهذا الرأي مقاس بالمقياس الترتيبي وليس هناك ثمة إثبات بأن المسافة بين التأييد التام والتأييد تعادل المسافة بين الرافض والرافض بشدة.

٣) المقياس الفتري (Interual Scale) :

هو قيماً ترتيبية رقمية وهو مقياس أعلى درجة من المقياس الترتيبي، فمع كونه ترتيبي سواء تصاعدياً أو تنازلياً إلا أن المسافات بين كل قيمـــة يأخـــذها المتغير المقاس بهذا المقياس وأي قيمة أخرى معلومة أي أن المتغير المقـــاس بهذا المقياس نستطيع معرفة موقعه ونستطيع معرفة الفرق بين القيمـــة التـــي يأخذها وبين أية قيمة أخرى على نفس المقياس فهو كالمقياس الأســـمي يعتبـــر تمييزاً بين الأشياء وأيضاً كالمقياس الترتيبي يعني أفضـــلية وترتيـــب ويزيـــد عليهما حساب الفترات بمعنى أن هناك مجال لتقدير المسافة بـــين كـــل قيمـــة والقيمة التي تليها ولكن الخاصية الوحيدة التي لا تتوفر في هذا المقياس والتـــي تميزه عن المقياس التالي هو نقطة الصفر فعلى هذا المقياس لا يكـــون هنـــاك معنى حقيقي للصفر وأوضح مثال لمتغير يقاس بهذا المقياس هو درجة حرارة الجو فأنه عندما تكون درجة الحرارة ٢٠ درجة هل يعني ذلك أن حرارة الجو تعادل ضعف حرارتها عندما تكون درجة الحرارة ١٠ درجات وتعادل نصـــف حرارتها عندما تكون درجة الحرارة ٤٠ بالطبع لا يجوز هذا التحليـــل ولكننـــا نعلم من هذا المقياس أنه عندما تكون درجة الحرارة ٢٠ درجة فإن الفرق يكون ٥ درجات بينها وبين درجة الحرارة ١٥ درجة أو بينها وبين درجة الحـــرارة ٢٥ وهكذا.

وهذا ما يميز هذا المقياس عن المقياس السابق ولكن القصـــور فـــي هـــذا المقياس يتمثل في تقدير النسب أي النسب في هذا المقياس ليســـت ذات معنـــى

فالمقياس الفتري عبارة عن قيم ترتيبية رقمية ولكن النسب بين القيم التي يأخذها المتغير المقاس بهذا المقياس ليست ذات معنى بسبب أن الصفر غيـــر حقيقـــي ووجود الصفر على هذا المقياس لا يعني فناء الشيء فعنـــدما تكـــون درجـــة الحرارة صفر لا يعني ذلك أنه لا توجد حرارة على الإطلاق.

٤) المقياس النسبي (Ratio Scale) :

هو أعلى درجات المقياس وهو غالباً المقياس المعروف لدينا وهـــو قيـــاس المتغيرات بقيم رقمية محسوبة ترتيبياً ونسبياً، فهو يميز بين الأشياء ويرتبهـــا ويحفظ المسافات بين قيمة المتغيرات ونسبتها لبعضها البعض والميزة التي تميز هذا المقياس عن سابقه هو نقطة الصفر الحقيقية التي تعني فناء الشيء، فالشيء الذي وزنه صفر لا يوجد والشيء الذي طوله صفر لا يوجد، وذلك يحدد نقطة بداية المقياس ويحفظ النسب بين القيم فعندما يكون وزن قطعة اللحـــم ٢ كيلـــو جرام فهي ضعف قطعة اللحم التي وزنها ١ كيلو جرام ونصف، ونصف قطعة اللحم التي وزنها ٤ كيلو جرام وهكذا.

وثوب القماش الذي طوله ٥٠ متر نصف طول ثوب القماش الـــذي طولـــه ١٠٠ متر وضعف طول ثوب القماش الذي طوله ٢٥ متر وهكذا.

وغالبية المقاييس المستخدمة هي من نوع المقياس النسبي فهي مثل الميزان الذي نزن به الأشياء أو المازورة التي تستخدم في قياس الأطوال.

على ذلك يتبين لنا أن قياس المتغير مشكلة تتوقف على طبيعة هذا المتغيـــر وأن المقياس المناسب سواء كان أسمي أو ترتيبي أو فتري أو نســـبي يتوقـــف استخدامه على نوع وطبيعة المتغيرات محل القياس.

الدالة (حيزها ومداها)

عندما نقول أن ص = د (س)

نعني بذلك أن ص دالة في س، بمعنى أن يمكن معرفة قيمة ص عندما نعرف قيمة س وعادة ما يكون ص هو المتغير التابع، س هو المتغير المستقل وجميع القيم التي يمكن أن يأخذها المتغير س تسمى حيز الدالة (Domain) فإن كانت القيم التي يمكن أن يأخذها المتغير س تتراوح بين ٥ أو ١٤ تقول أن حيز الدالة من ٥ – ١٤ وإذا كانت س ممكن أن تأخذ أي قيمة سواء كانت سالبة أو موجبة قلنا أن حيز الدالة من –٥٥ إلى +٥٥.

وجميع القيم التي يمكن أن يأخذها المتغير ص تسمى مدى الدالة (Range)، فإن كانت قيم ص بعد التعويض بقيمة س نأخذ القيمة من ٢٠ إلى ٤٠ مثلاً نقول أن مدى الدالة من ٢٠ – ٤٠ مثال ذلك إذا كانت العلاقة بين طول تلاميذ المرحلة الابتدائية وأعمارهم كالتالي : ص =١٠٠ + ٣ س حيث س يرمز إلى عمر التلميذ، ص يرمز إلى طوله فمن المعروف أن عمر تلميذ المرحلة الابتدائية يتراوح بين خمس سنوات إلى ١٢ سنة وبالتالي فقيمة المتغير س تتراوح بين ٥ ، ١٣ وهو ما نطلق عليه حيز الدالة.

فإذا ما عوضنا في المعادلة المذكورة عن قيمة س بـ ٥ ثم عوضنا عن س بـ ١٣ ينتج لنا أن قيمة ص ستكون بين ١١٥، ١٣٩ على الترتيب وهذا ما نطلق عليه المدى حيث يكون مدى الدالة ص = د (س) من ١١٥ إلى ١٣٩.

وجدير بالذكر أنه أحياناً يطلق على حيز الدالة نطاقها فتستخدم كلمتي حيـز ونطاق بمعنى واحد وهو جميع القيم التي يمكن أن يأخذها المتغير المسـتقل أو المتغير الداخل في حساب قيمة الدالة أما المدى فكما قلنا فهو القيمة التي يمكـن أن تأخذها الدالة نفسها.

الفصل الثالث

جمع البيانات الإحصائية

Statistical Data Collection

تعد مرحلة جمع البيانات الإحصائية من أهم مراحل البحث الإحصائي، حيث أن توافر البيانات الإحصائية بالقدر والكيف المناسبين لدراسة المشكلة البحثية يعتبر من المقومات الأساسية لدقة ومصداقية النتائج المتحصل عليها من التحليل الإحصائي، ومن ثم إمكانية التوصل إلى القرار السليم بشأن حل تلك المشكلة.

نوعية البيانات الإحصائية :

يمكن تقسيم البيانات الإحصائية وفقاً لنوعيتها إلى :

١) **بيانات وصفية** : هي تلك البيانات التي تمثل أوصافاً معينة ولا يمكن التعبير عنها في صورة رقمية، ومنها على سبيل المثال بيانات الجنس (ذكر – أنثى)، الجنسية (عربي – أجنبي)، الحالة الاجتماعية (أعزب – متزوج – أرمل – مطلق) ... الخ.

٢) **بيانات كمية** : وهي تلك البيانات التي يمكن التعبير عنها في صورة رقمية، وهذه يمكن تصنيفها إلى :

أ) **بيانات عددية** : تتمثل في كافة البيانات التي يمكن عدها ومنها على سبيل المثال الإحصاءات السكانية، عدد السائحين في موسم معين، عدد الفنادق السياحية في مدينة القاهرة ... الخ.

ب) **بيانات مقيسة** : وهي تلك البيانات التي يمكن الحصول عليها عن طريق القياس مثل الأوزان، الأطوال ودرجة الحرارة.

مصادر البيانات الإحصائية :

يمكن تصنيف البيانات الإحصائية وفقاً لمصدر الحصول عليها إلى :

١) **بيانات أوليـــة :** هي تلك البيانات التي يتم تجميعها لأول مرة من مصادرها الأساسية أو ما تعرف بالبيانات الميدانية. وتتميز هذه البيانـــات بدقتها وملاءمتها لهدف البحث الذي جمعت من أجله، غير أنها تتطلب جهداً شاقاً وتكلفة عالية نسبياً.

٢) **بيانات ثانوية :** وهي تلك البيانات المنشورة أو غير المنشـــورة التـــي تـــم جمعها بمعرفة الهيئات أو المؤسســـات الحكوميـــة أو بـــاحثين آخرين في فترات زمنية سابقة على إجراء البحث وذات علاقة وثيقة بالمشكلة موضع التحليل. ويتسم هذا المصـــدر بســـهولة وسرعة الحصول على البيانات المطلوبة وبتكلفة أقل.

هذا وتفضل البيانات الأولية عن البيانات الثانوية للأسباب التالية :

- تتيح البيانات الأولية للباحث تفاصيل أكثر وأدق عن المشكلة البحثية موضع التحليل بالمقارنة بالبيانات الثانوية.
- زيادة درجة ثقة الباحث في البيانات الأولية عن نظيرتها الثانويـــة نظـــراً لمساهمته في جمعها.
- عدم وضوح بعض المفاهيم والتعاريف في البيانات الثانوية، هذا فضلاً عن احتمال تباينها أو تعارضها بين مصادرها المختلفة.
- ارتفاع درجة احتمال الخطأ عند نقل أو نشر البيانات الثانوية.

الأخطاء الشائعة في جمع البيانات :

تتمثل أهم الأخطاء الشائعة في جمع البيانات فيما يلي :

١) التحيز : يعتبر خطأ التحيز أحد الأخطاء الشائعة عند جمع البيانات وقد يكون هذا التحيز مقصوداً كأن يقوم جامع البيانات بتوجيه أسئلة الاستبيان بطريقة توحي للمبحوث بإجابة معينة، كما قد يكون التحيز لا شعورياً بمعنى أن الباحث يكون متأثراً باتجاهاته وميوله الشخصية في تفسيره للحقائق المحيطة به.

٢) إغفال بعض المتغيرات الهامة : أحياناً ما يغفل الباحث بعض المتغيرات التي قد يكون ذات ارتباط قوي بالمشكلة موضع الدراسة.

٣) إهمال عملية مراجعة استمارات الاستبيان : الأمر الذي يترتب عليه عدم دقة العمليات الحسابية ومن ثم عدم صحة البيانات التي تم جمعها.

٤) عدم مراعاة التناسب بين مفردات العينة المسحوبة من المجتمع : الأمر الذي قد يصعب معه إجراء المقارنة بين بعض مفرداتها.

٥) محدودية البيانات المجموعة وعدم كفايتها لتحليل جوانب المشكلة البحثية وبالتالي عدم التأكد من مصداقية القرارات المعتمدة عليها.

٦) عدم وضوح المفاهيم الواردة باستمارة الاستبيان بالنسبة لجامعي البيانات، مما يؤدي إلى عدم دقة النتائج المتحصل عليها من تحليل مثل هذه البيانات.

أساليب جمع البيانات :

يمكن جمع البيانات الإحصائية بأي من الأسلوبين التاليين :

أ) أسلوب التعداد أو الحصر الشامل Census :

ويقصد بهذا الأسلوب جمع البيانات الإحصائية المطلوبة من كل مفرده من مفردات المجتمع المحدد طبقاً للمشكلة موضع البحث، وذلك دون ترك أي مفردة من مفردات المجتمع دون جمع بيانات عنها ومثال ذلك التعداد الشامل للسكان أو التعداد الزراعي العام.

ومن مزايا هذا الأسلوب أنه يعطي نتائج دقيقة إذا ما تم فحص كل مفردة من مفردات المجتمع فحصاً دقيقاً مع تسجيل نتائجه بدقة.

أما عيوب هذا الأسلوب فأنها تتمثل فيما يلي :

١) صعوبة جمع بيانات كافية عن كل مفردة من مفردات المجتمع خاصة مع كبر عدد هذه المفردات.

٢) استحالة الحصر الشامل لمفردات المجتمعات غير المحدودة مثل مجتمع الأسماك في البحار.

٣) قد يستلزم إتباع هذا الأسلوب وقتاً طويلاً نسبياً قد يستحيل معه إجراء البحث، كما قد يكون مستحيلاً من الناحية الاقتصادية.

٤) لا يمكن إتباع هذا الأسلوب في حالة الفحص المدمر Destructive Testing أي عندما يتسبب فحص الوحدات في تدميرها.

ب) أسلوب المعاينة Sampling :

يعتمد هذا الأسلوب على جمع بيانات عن المشكلة موضع البحث من مجموعة مختارة بطريقة معينة من مفردات البحث تمثل جزءاً صغيراً نسبياً من الحجم الكلي للمجتمع الخاص بالمشكلة ثم يستخلص من بيانات هذه المفردات نتائج تصلح للتعبير عن المجتمع بأكمله، أي أنها تعميم من جزء إلى الكل وبشرط أن تكون العينة المختارة ممثلة تمثيلاً كاملاً وصادقاً للمجتمع الأصلي.

العوامل المحددة للطريقة المناسبة لجمع البيانات الإحصائية :

يتوقف اختيار الأسلوب المناسب لجمع البيانات على العوامل التالية :

أ) طبيعة المجتمع : ويقصد به المجتمع الإحصائي موضع الدراسة Population أي مجموعة المفردات التي تتصف ببعض الخصائص المشتركة العامة والتي يمكن من خلالها التعرف عليها وقد يكون المجتمع محدداً مثل الإنتاج الكلي من محصول معين أو المساحة المنزرعة بمحصول ما ... الخ، أو يكون المجتمع غير محدد مثل عدد الأسماك في بحيرة ما. ومن الملاحظ أن هناك العديد من المجتمعات المحدودة غير أنها تتعامل كأنها غير محدودة كعدد ثمار البرتقال في مزرعة معينة.

ويمكن القول بأن المفاضلة بين أسلوبي الحصر الشامل والعينة تتم عندما يكون المجتمع موجوداً بالكامل ويمكن التعرف على جميع مفرداته أما في بقية الحالات فإن أسلوب المعاينة هو الأسلوب الوحيد الذي يمكن استخدامه للحصول على البيانات الإحصائية عن المجتمع موضع

الدراسة. غير أن هناك بعض الحالات التي تتسم فيها مفردات المجتمع بالتقلبات في خصائصها كما هو الحال في الدراسات الخاصة بسلوك المستهلك، مما يجعل من المفضل استخدام أسلوب العينة، حيث أن إجراء الحصر الشامل يتطلب وقتاً طويلاً نسبياً قد يحدث خلاله تغيراً في النمط الاستهلاكي ومن ثم تصبح البيانات الإحصائية مجافية للواقع.

ب) طيعة البيانات المطلوبة : يتطلب جمع البيانات الإحصائية في بعض الحالات تدمير مفردات المجتمع موضع الدراسة أو إتلافها كما في حالة دراسة عدد كرات الدم فلا يعقل إجراء العد في كل الدم، بل يكتفي بعينة منه. كذلك في حالة دراسة القوى التدميرية للقنابل التي ينتجها منع حربي معين، فإنه من غير المنطقي أن تفجر جميع القنابل المنتجة بل تتم الدراسة على عينة منها، كما قد تكون البيانات الإحصائية متشعبة وكثيرة بحيث تتطلب وقتاً طويلاً لجمعها ودقة وخبرة من الباحث للحصول عليها، فعندئذ يكون أسلوب المعاينة هو الأنسب، حيث تكون الدراسة على نطاق ضيق يمكن معه توخي الدقة في جمع البيانات، ومن أمثلة هذا النوع من الدراسات تلك الخاصة بنفقات المعيشة والتي تتطلب جمع البيانات عن أوجه الإنفاق المختلفة من سلع وخدمات.

ج) الإمكانيات المادية والفنية والفترة الزمنية المخصصة للدراسة : تعتبر هذه العوامل من بين العوامل المحددة للاختيار ما بين أسلوبي الحصر الشامل والمعاينة. ومن الطبيعي أن أسلوب المعاينة أقل تكلفة ويتطلب عدداً أقل من الفنيين ويستغرق وقتاً أقصر نسبياً لإجرائه.

خطأ المعاينة
Sampling Error

يقصد بخطأ المعاينة ذلك الخطأ الناشئ عن عدم اختيار جميع مفردات المجتمع والاكتفاء باختيار البعض منها كعينة، مما يؤدي إلى ظهور فرق بين النتائج التي يمكن الحصول عليها من العينة وتلك الخاصة بالمجتمع تحت نفس الظروف. وترجع الفروق الناتجة من بيانات العينات المسحوبة من المجتمع إلى الصدفة أو ما يسمى بخطأ المعاينة والذي يمكن تقدير قيمته بدرجات معينة من الدقة وبإتباع الطرق والأساليب الإحصائية. ويتوقف خطأ المعاينة على العوامل التالية :

١) **حجم العينة** : فكلما ازداد حجم العينة كلما قل خطأ المعاينة وتزايدت الثقة في النتائج المتحصل عليها.

٢) **طريقة اختيار العينة** : حيث أنه من المنطقي أن اختيار الطريقة المناسبة العملية للمعاينة يقلل من خطأ المعاينة والعكس صحيح.

٣) **تباين المجتمع** : من الثابت أن خطأ المعاينة يقل كلما كانت مفردات المجتمع أكثر تجانساً فيما يتعلق بالظاهرة موضع الدراسة ومن المفضل زيادة حجم العينة عندما يكون المجتمع أقل تجانساً وتقليل حجمها عندما يكون المجتمع أكثر تجانساً.

وتنقسم العينات من حيث الحجم إلى عينات صغيرة وعينات كبيرة وتتكون العينة الصغيرة من مائة وحدة فأقل، ويجب الاهتمام بدرجات الحرية لأنها تؤثر

على المقاييس المستخدمة تأثيراً ملموساً في تلك العينة أما العينات الكبيرة فهي التي يزيد عدد وحداتها عن ١٠٠ ويضطر الباحث إلى تبويب قيم هذه الوحدات على شكل توزيع تكراري وتوزيع احتمالي نظراً لكثرة عددها. ويمكن في هذه العينات الكبيرة تجاهل درجات الحرية نظراً لضعف تأثيرها على المقاييس المستخدمة.

أساليب المعاينة الإحصائية
Statistical Sampling Techniques

هناك العديد من الأساليب الإحصائية المستخدمة في إجراء عملية المعاينة وذلك للحصول على عينات ممثلة للمجتمع الأصلي موضع الدراسة وسنتناول هنا أهم الأساليب الشائعة للمعاينة الإحصائية.

أولاً : المعاينة العشوائية البسيطة Simple Random Sampling :

يقصد بالعينة العشوائية البسيطة تلك العينة التي تعطي كل مفردة على مفردات المجتمع نفس الفرصة للظهور فيها.

A Random Sample is one in which each individual in the population is equally likely to be sampled.

وعلى ذلك يجب أن يكون اختيار العينة من شأنه ضمان الفرصة المتكافئة، حيث يتطلب إختيار العينة العشوائية البسيطة من مجمع توافر الشرطين التاليين:

أ) أن تكون جميع مفردات المجتمع الأصلي متجانسة من حيث الظاهرة موضع الدراسة.

ب) أن يكون احتمال اختيار أي مفردة من مفردات المجتمع في العينة متساوياً فلو كان عدد مفردات المجتمع الأصلي هو N كان احتمال ظهور أي مفردة من مفردات المجتمع في العينة المختارة هو :

$$P = \frac{1}{N}$$

وفيما يلي أساليب اختبار العينة العشوائية البسيطة :

١) تخصيص بطاقة لكل مفردة من مفردات مجتمع البحث وبعد ذلك تخلط البطاقات خلطاً جيداً ثم يسحب من بينها عدد من البطاقات بطريقة عشوائية بحيث يكون عددها مماثلاً لعدد المفردات المطلوبة للعينة وتكون الأرقام الموجودة على البطاقات المسحوبة هي أرقام المفردات التي تتكون منها العينة.

٢) استخدام كرات متشابهة الشكل والحجم والوزن يبلغ عددها مجموع مفردات مجتمع البحث وتحمل أرقامه أي على كرة رقم. وبعد أن تخلط الكرات خلطاً جيداً يتم الاختيار العشوائي لكرات عددها حجم العينة المطلوبة. وعلى ذلك تتكون العينة من مفردات المجتمع التي تحمل الأرقام المقابلة لتلك المبينة على الكرات المختارة وتجدر الإشارة إلى أن هاتين الوسيلتين لا تصلحان في اختيار العينة عندما يكون عدد مفردات المجتمع كبيراً جداً وهذا هو الشائع. ولذلك يستعين الباحث بجداول الأعداد العشوائية.

٣) استخدام جداول الأعداد العشوائية Table of Random Number ويتكون من قوائم من الأرقام جاء ترتيبها على أساس من الصدفة وذلك وفقاً لما أثبتته كافة الاختبارات الإحصائية والعملية الممكنة وعند سحب العينة باستخدام جدول الأعداد العشوائية تتبع الخطوات التالية :

أ) يخصص لكل مفردة من مفردات مجتمع البحث رقم مسلسل من ١ إلى N (مجموع المفردات).

ب) تتخذ عن طريق الصدفة نقطة بدء على جدول الأعداد الممثلة للمفردات التي تتكون منها العينة المطلوبة، ويكون اختيار الأعداد أفقياً أو رأسيا على التوالي. وعندما يكون مجتمع البحث شاملاً لعشر مفردات أو أقل فإن كل رقم من الأرقام الممثلة لمفردات العينة يتكون من عدد واحد ويخصص الصفر للمفردة التي تحمل الرقم (١٠)، ويتكون من عددين إذا بلغ عدد مفردات مجتمع البحث ما بين ١١ ، ١٠٠ مفردة مع تمييز المفردة التي يخصها الرقم ١٠٠ بصفرين، ويتكون من ثلاث أعداد إذا بلغ عدد مفردات مجتمع البحث ما بين ١٠١، ١٠٠٠ مفردة مع تمييز المفردة التي يخصها الرقم ١٠٠٠ بثلاثة أصفار وهكذا.

ج) يستبعد تكرار الأرقام التي سبق اختيارها وكذلك يهمل كل رقم يزيد عن مجموع مفردات مجتمع البحث.

مثال :

إذا أردنا اختيار ٧ أرقام عشوائية من ٧٠٠ فرد فتعطى هذه الأفراد أرقام مسلسلة ابتداء من ٠٠١ إلى ٧٠٠ ثم نعتبر أن كل عدد مكون من ثلاثة أرقام، تقرأ الأعداد العشوائية من ثلاثة أعمدة ولتكن الثلاثة الأولى من اليمين من الجدول المرفق فتكون المفردات هي :

٣٤٢، ٠٣٢، ٣٠٥، ٢٣١، ٤٧٧، ٠٩٨، ٢٥٢

ويلاحظ أننا أهملنا العدد الأول ٨٤٦ والعدد ٧١٧، حيث أن قائمة المفردات لا تشتمل هذه الأعداد.

وتجدر الإشارة إلى أن العينة العشوائية البسيطة قد يطلق عليها العينة المطلقة Unrestricted Random Sample وفي هذا النوع من العينات العشوائية يكون هناك فرصة متكافئة أمام أي مجموعة (n) من الوحدات التي تكون حجم العينة للاختيار فمثلاً إذا كان المجتمع المطلوب دراسته يتكون من ٢٥ مزارعاً ويراد أخذ عينة منه لدراسة ظاهرة معينة وكان حجم هذه العينة يساوي ٥ مزارعين فقط فهناك عدد من الطرق التي يمكن بواسطتها اختيار هذه العينة يساوي :

عدد مفردات المجتمع = N $\qquad {}_N C_n = \dfrac{Ni}{n!(N-n)!}$

حجم العينة = n

$${}_{25}C_5 = \frac{25!}{5!(25-5)!} = 53130$$

توفيقية C = Combination

جدول أرقام عشوائية

٤٦	٩٨	٦٣	٧١	٦٧	٣٣	٢٦	١٦	٠٨	٤٥
٤٢	٥٣	٣٢	٣٧	٣٢	٢٧	٠٧	٣٦	٠٧	٥١
٣٢	٩	٧٩	٧٨	٥٣	١٣	٥٥	٣٨	٥٨	٥٩
٠٥	٠٣	٧٢	٩٣	١٥	٥٧	١٢	١٠	١٤	١٢
٣١	٦٢	٤٣	٠٩	٩٠	٠٦	١٨	٤٤	٣٢	٥٣
١٧	٣٧	٩٣	٢٣	٧٨	٨٧	٣٥	٢٠	٩٦	٤٣
٧٧	٠٤	٧٤	٤٧	٦٧	٢١	٦٧	٣٣	٥٠	٢٥
٩٨	١٠	٥٠	٧١	٧٥	١٢	٨٦	٧٣	٥٨	٠٧
٥٢	٤٢	٠٧	٤٤	٣٨	١٥	٥١	٠٠	١٣	٤٢

٤٩	١٧	٤٦	٠٩	٦٢	٩٠	٥٢	٨٤	٧٧	٢٧
٧٩	٨٣	٨٦	١٩	٦٢	٠٦	٧٦	٥٠	٠٣	١٠
٨٣	١١	٤٦	٣٢	٢٤	٢٠	١٤	٨٥	٨٨	٤٥
٠٧	٤٥	٣٢	١٤	٠٨	٣٢	٩٨	٩٣	٠٧	٧٢
٠٠	٥٦	٧٦	٣١	٣٨	٨٠	٢٢	٠٢	٥٣	٥٣
٤٢	٣٤	٠٧	٩٦	٨٨	٤٥	٤٢	٠٦	٨٧	٩٨
١٣	٨٩	٥١	٠٣	٧٤	١٧	٧٦	٣٧	١٣	٠٤
٩٧	١٢	٢٥	٩٣	٤٧	٧٠	٣٣	٢٤	٠٣	٥٤
١٦	٦٤	٣٦	١٦	٠٠	٠٤	٤٣	١٨	٦٦	٧٩
٤٥	٥٩	٣٤	٦٨	٤٩	١٢	٧٢	٠٧	٣٤	٤٥
٢٠	١٥	٣٧	٠٠	٤٩	٥٢	٨٥	٦٦	٦٠	٤٤

ثانياً : المعاينة المنتظمة Systematic Sampling :

وهي من أهم الطرق الشائعة في اختيار العينات وفكرتها الأساسية هي تساوي الفترات بين الوحدات المختارة في العينة. فإذا فرضنا أننا نريد أخذ عينة حجمها ١٠ أفراد من مجتمع عدد أفراده يساوي ١٠٠ فإن طريقة أخذ عينة منتظمة تكون كما يلي:

١) تحدد الفترة بين الوحدات وهي تساوي :

$\frac{N}{n}$ حيث N = عدد أفراد المجتمع.

n = عدد أفراد العينة

$$= \frac{١٠٠}{١٠} = ١٠$$

٢) تختار الوحدة الأولى من وحدات العينة عشوائياً من الفتـرة الأولـى فـي الإطار وليكن الوحدة رقم (٦) مثلاً.

٣) تكون أرقام الوحدات المختارة في العينة هي :

٦ ، ١٦ ، ٢٦ ، ٣٦ ، .... ، ٩٦.

وليس من المهم معرفة عدد أفراد المجتمع على وجه الدقة لتحديد الفترة بين الوحدات، وتعد هذه الطريقة في سحب العينات مرضية إذا لم يكن ثمة تغيرات منتظمة في الإطار، فإذا وجدت مثل هذه التغيرات فإن العينة سوف تعطي نتائج متحيزة إلى حد بعيد. وتتميز هذه الطريقة بسهولتها وصدق تمثيلها للعينة بشكل منتظم إذ تتوزع العينة على إطار المجتمع توزيعاً منتظماً الأمر الذي لا يضمنه اختيار عينة عشوائية بسيطة. إلا أنه من ناحية أخرى فإنه طالما أن نتائج العينة تتأثر بترتيب وحدات المجتمع في الإطار فليس من الممكـن معالجـة أخطـاء المعاينة رياضياً على نحو مقبول كما هو الحال عليه فـي العينـة العشـوائية البسيطة. وهذه المعاينة تجمع ما بين العشوائية والانتظام.

ثالثاً : المعاينة العشوائية الطبقية Stratified Random Sampling :

وتعتبر هذه الطريقة تطويراً للعينة العشوائية البسيطة عندما يمكـن تقسـيم وحدات المجتمع إلى طبقات Starts متجانسة ومحددة الخصائص وطبقاً لهـذه الطريقة يقسم المجتمع إلى مجموعات وتختار الوحدات من بين هذه المجموعات طبقاً لخطة المعاينة، فإذا كانت هذه الخطة تناسبية Proportional فإن معنـى

ذلك أن عدد الوحدات التي يتم اختيارها من كل مجموعة تتناسب مع حجم هذه المجموعة في المجتمع وذلك كما يلي :

حجم العينة = n

حجم الطبقة = n

حجم المجتمع = N

∴ الأهمية النسبية للطبقة في المجتمع = $\frac{n_s}{N} 100$

وعلى ذلك يمكن إيجاد عدد الوحدات التي يلزم اختيارها من الطبقة الأولى (n_1) كالآتي :

$$n_1 = \frac{n_{s1}}{N}.n.$$

ويلاحظ أن هذه الطريقة تنطوي على اختيار عشوائي لكافة أفراد العينة بالنسبة لكل طبقة من الطبقات، ومن ثم فإن هناك فرصة متساوية لكل فرد من أفراد العينة في الاختيار، ومن ثم فإنه يمكن النظر إلى العينة الطبقية على أنها عينة عشوائية مقيدة، وتمتاز هذه الطريقة بدقة تمثيل العينة للمجتمع المأخوذة منه نتيجة لتضمين العينة وحدات من كافة طبقات المجتمع الأمر الذي لا يمكن ضمانه في الحالتين السابقتين.

وقد يفضل أحياناً اختيار العينة الطبقية بإتباع توزيع غير تناسبي Disproportionate Allocation إذا كانت بعض الطبقات تتميز بثباتها النسبي في حين أن معدل التغير في الطبقات الأخرى أعلى بكثير، فإنه يلزم التركيز على الطبقات التي تظهر فيها تغير واضح وإعطائها وزن أكبر من حجمها النسبي في المجتمع، لأن الطبقات التي ليس بها تغير يذكر يكفي لتمثيلها عدد قليل من الأفراد دون أي إخلال بصحة النتائج.

مثال : المطلوب اختيار عينة من ١٠٠٠ سائح من مجتمع يحتوي على أربع طبقات وذلك على النحو التالي :

جدول (١) : توزيع السائحين الوافدين إلى مصر حسب جنسياتهم عام ١٩٩٢ :

الطبقة	الجنسية	عدد السائحين بالألف	الأهمية النسبية
الأولى	عرب	١١٠٣	%٣٤
الثانية	أوربيون	١٥٥٥	%٤٩
الثالثة	أمريكان	٢٢٤	%٧
الرابعة	آخرون	٣٢٥	%١٠
الإجمالي		٣٢٠٧	١٠٠

حجم العينة المسحوبة من الطبقة الأولى : $n_1 = \frac{34}{100} \times 1000 = 340$

حجم العينة المسحوبة من الطبقة الثانية : $n_2 = \frac{49}{100} \times 1000 = 490$

حجم العينة المسحوبة من الطبقة الثالثة : $n_3 = \frac{7}{100} \times 1000 = 70$

حجم العينة المسحوبة من الطبقة الرابعة : $n_4 = \frac{10}{100} \times 1000 = 100$

أما في حالة ما إذا كان التوزيع غير تناسبي، فإنه يتم إعطاء أوزان للطبقات الأربع (W) وذلك على النحو التالي :

الجنسية	عدد السائحين بالألف n_s	الانحراف المعياري W	$n_s W$	الأهمية النسبية
عرب	١١٠٣	١	١١٠٣	١٦%
أوروبيون	١٥٥٥	٣	٤٦٦٥	٦٧%
أمريكان	٢٢٤	١	٢٢٤	٣%
آخرون	٣٢٥	٣	٩٧٥	١٤%
الإجمالي	٣٢٠٧		٦٩٦٧	

$$n_1 = \frac{16}{100} \times 1000 = 160$$

$$n_2 = \frac{67}{100} \times 1000 = 670$$

$$n_3 = \frac{3}{100} \times 1000 = 30$$

$$n_4 = \frac{14}{100} \times 1000 = 140$$

وبالتالي تتوزع العينة على النحو التالي :

الطبقة	جنسية السائحين	حجم العينة المسحوبة
الأولى	عرب	١٦٠
الثانية	أوربيون	٦٧٠
الثالثة	أمريكان	٣٠
الرابعة	آخرون	١٤٠
الإجمـــــــالي		١٠٠٠

رابعاً : المعاينة متعددة المراحل Multistage Random Sampling :

يستخدم هذا الأسلوب في كثير من الدراسات التي تتطلب اســتخدام بعــض أساليب المعاينة بحيث تمر بعده مراحل للوصول إلى المفردات الأوليــة محــل الدراسة. فمثلاً لدراسة عينة من الأسر للتعرف على مستواهم المعيشي، فإنــه يلزم تقسيم الجمهورية إلى محافظات ونختار عينة ممثلة من هذه المحافظــات وذلك من خلال تطبيق أي أسلوب من الأساليب السابقة بحيث تكون ممثلة تمثيلاً صادقاً لجميع أنحاء الجمهورية. وهذه هي المرحلة الأولى، أما فــي المرحلــة الثانية فسيتم اختيار عينة عشوائية من الوحدات الإدارية أو المراكــز الممثلــة للمحافظات المختارة في المرحلة الأولى، والمعروف أن كل مركز يضم عــدد من القرى، وعلى ذلك فإن المرحلة التالية هي اختيار عدد من القرى من كــل مركز وقد يتم ذلك عشوائياً أو بأخذ عينة طبقية بعد تقسيم القرى في المركــز إلى طبقات (قرى كبيرة ــ صغيرة ــ نجوع ... الخ) والخطوة الأخيــرة هــي

اختيار عينة منتظمة من المنازل داخل القرى التي وقع عليها الاختيار في المرحلة السابقة. ومن الملاحظ أن هذا التصميم ليس هو التصميم الوحيد فقد يمكن أخذ عينة منتظمة أو طبقية في المرحلة الأولى أو عشوائية في الثانية وهكذا ... ويتوقف ذلك على طبيعة المشكلة موضع الدراسة والغرض من البحث ونوع الوحدات.

الاختيار غير العشوائي للعينات

كثيراً ما تستعمل في الدراسات الاقتصادية عينات لم تسحب بالطرق العشوائية سالفة الذكر إما لأن الاختيار غير العشوائي قد يعطي أحياناً نتائج أفضل أو لأن اختيار العينة عشوائياً قد يستغرق وقتاً طويلاً وتكاليف كبيرة. وفيما يلي بعض هذه الأساليب :

١) المعاينة العمدية Purposive Sampling :

تظهر أهمية هذا الأسلوب عندما يتكون المجتمع من وحدات قليلة نسبياً معروفة الخصائص لهؤلاء الذين يصممون خطة العينة. ومن ثم يمكن اختيار وحدات العينة على النحو الذي يضمن توافر مجموعة من خصائص المجتمع الهامة ذات الصلة بموضوع الدراسة. فإذا سحبت عينة عشوائية من مصانع الأغذية المحفوظة وتبين أنها بالصدفة لا تضم شركة قها وهي من كبريات شركات الأغذية المحفوظة سوف تكون محدودة القيمة من حيث استخلاص نتائج معقولة عن صناعة الأغذية المحفوظة. في هذه الحالة تعطي المعاينة العمدية نتائج أفضل بكثير. أما حيث يزيد عدد أفراد المجتمع ويضم وحدات كثيرة صغيرة كالأسر، الأفراد المصانع الصغيرة وغيرها فإن الاختيار العمدي

يصبح عديم القيمة لافتقاره إلى أساس الاختيار وهو معرفة الخصائص الهامة التي ترغب في أن تتضمنها العينة. ومع أن طريقة الاختيار العمدي للعينات تستعمل بكثرة في الدراسات الاقتصادية وخاصة عندما يكون الغرض من الدراسة اتخاذ قرار بشأن موضوع معين وبالتالي لا يكون هناك وقت كاف لاختيار عينة بطرق أخرى، فإنه ينبغي التنويه بأن مثل هذه العينات قد تعطي نتائج مضللة إذا لم تكن الضوابط المستخدمة على درجة عالية من حسن تمثيل المجتمع وإذا تدخلت عوامل أخرى في اختيار أفراد العينة كالقرب من مكان الدراسة أو سهولة المواصلات ... الخ الأمر الذي يؤثر على موضوعية الاختيار.

٢) المعاينة الحصصية Quota Sampling :

تعتبر المعاينة الحصصية إحدى الوسائل الشائعة لجمع البيانات عن طريق المقابلة الشخصية وتتميز هذه الطريقة بأن جامعي البيانات الذين يقومون بهذه المقابلات الشخصية Interviews يترك لهم كذلك اختيار الأفراد في العينة إذ يكتفي بإعطائهم تعليمات بعدد المقابلات التي يقومون بها في منطقة معينة وتوزيع هذه المقابلات بين الفئات أو المجموعات المختلفة أو وفقاً لأي تصنيف تقسم المنطقة إلى أجزاء أو حصص. على أن هذه الطريقة في المعاينة تتصف بسهولة وقوعها في خطأ التحيز نتيجة لترك الحرية لجامعي البيانات في اختيار أفراد العينة وبالتالي اتجاههم إلى اختيار المعارف مثلا أو الوحدات التي يسهل الوصول إليها أو غير ذلك، وفضلاً عن ذلك فإن هذه الطريقة تفتقر إلى الضوابط التي تتوافر في المعاينة العمدية.

٣) المعاينة السهلة Convenience Sampling :

ويقصد بهذا الأسلوب من أساليب اختيار العينات كل الطرق التي تستهدف اختيار عينة على أساس ملائمة هذا الاختيار للباحث أو لجامعي البيانات. فقد يلجأ باحث إلى توزيع صحيفة الاستبيان على أصدقائه باعتبار أنهم يمثلون عينة عشوائية ويحاول دراسة توزيع اتفاقهم الاستهلاكي الشهري. ومن الواضح أن مثل هذه العينة سوف تعطي نتائج مضللة إلى حد بعيد لأن هذه المجموعة من الأصدقاء تمثل عدداً من الأفراد تجمعهم صفات مشتركة مثل تقارب السن والدخول والحالة الاجتماعية وغير ذلك وبالتالي فإن العينة لا يمكن أن تمثل مجتمع المستهلكين.

الفصل الرابع
عرض البيانات

عملية عرض البيانات الهدف منها التسهيل على القارئ بإعطائه صورة سريعة ومبسطة عن البيانات محل البحث يمكنه من خلالها تتبعها ومقارنتها وأخذ فكرة سريعة عن العلاقات الموجودة بينها.

وينقسم عرض البيانات إلى قسمين :

١) العرض الجدولي (التوزيعات التكرارية).

٢) العرض البياني (الخط البياني والرسوم البيانية)

أولاً : العرض الجدولي

يتم عرض البيانات في جداول تكرارية، والجدول التكراري يتكون من عمودين العمود الأول للفئات أو لأوجه الظاهرة، والعمود الثاني يبين تكرار الفئة أو تكرار الظاهرة والتكرار هو عدد مرات حدوث الظاهرة. والجدول بهذا المعنى يطلق عليه التوزيع التكراري، فالتوزيع التكراري وسيلة لتصنيف البيانات التي سبق جمعها والباحث هنا يقوم بدور فراز البريد الذي يفرز الخطابات حسب الجهة المرسلة إليها. لأن الباحث هنا هو الذي يقوم بتحديد الفئات وتقسيمها تقسيماً يسهل إدراك ما بينهما من علاقات ويوضح صفاتها ودلالتها.

وأوضح مثال على ذلك إذا كان لدينا بيانات عن نتائج اختبار عدد ١٠٠ طالب مثلاً وكان درجات الطلاب النهائية كالتالي ٥٤، ٧٦، ٦٣، ٨٢، ٩١،

٤٤، ... وهكذا مائة بيان للمائة طالب (القيم المعطاة). كما قلنا سابقاً أن الأرقام بهذه الصورة المبعثرة لا يمكن الاستفادة منها ولا يمكن بسهولة معرفة أعداد الناجحين أو أعداد الراسبين أو عدد الحاصلين على تقدير معين.

ولكن إذا جدولنا هذه البيانات في صورة توزيع تكراري فإنه كما سنرى يسهل علينا معرفة ما تبقى ونستخلص من البيانات أقصى ما يمكن أن تعطيه.

التوزيع التكراري

لعمل التوزيع التكراري ينبغي أن نحدد الحدين الأدنى والأعلى للقيم المعطاه ـ نحدد بعد ذلك الفئات بينهما، وعملية تحديد الفائت أمر متروك للباحث حيث أنه لا توجد قاعدة قاطعة مانعة لتحديد عدد الفئات أو لتحديد طول كل فئة ولكنه أمر تمليه طبيعة البيانات محل البحث وإن كان يفضل دائماً أن تكون الفئات متساوية وأن يكون عدد الفئات بين خمسة وعشرة فئات ولكن طبيعة البيانات والهدف من الجدولة له اليد العليا في ذلك.

وبالنسبة لمثالنا عن المائة طالب ودرجاتهم فإن تقسيم الطلبة إلى فئات توضح التقديرات التي حصلوا عليها تملى علينا أحد التقسيمين التاليين وذلك طبعاً حسب النظام المعمول به في الكلية أو المعهد محل الدراسة.

ضعيف جداً	٠ – ٣٥			
ضعيف	٣٥ – ٥٠	(F)	راسب	٠ – ٦٠
مقبول	٥٠ – ٦٥	(D)	مقبول	٦٠ – ٧٠
جيد	٦٥ – ٧٥	(C)	جيد	٧٠ – ٨٠
جيد جداً	٧٥ – ٩٠	(B)	جيد جداً	٨٠ – ٩٠
ممتاز	٩٠ – ١٠٠	(A)	ممتاز	٩٠ – ١٠٠

ومن هذين التقسيمين يتضح لنا أن أحدهما يتكون من ٦ فئات بينما الآخر يتكون من خمس فئات فقط وذلك دليل على أنه ليس هناك قاعدة محددة لتحديد عدد الفئات ولا طول كل فئة ولكن يراعي أن تشتمل أول فئة أصغر البيانات أو الحد الأدنى للقيم وأن تشتمل أخر فئة على أكبر البيانات أو الحد الأعلى للقيم المعطاة، وجدير بالذكر بأن الشرطة (-) الموجودة بين الحد الأعلى والحد الأدنى للفئة لها تعريف دولي وهو "أقل من". وهذا التعريف يفك الاشتباك بين الفئة والفئة التي تليها ويزيل اللبس والحيرة عند وضع القيم داخل الفئات فإذا قلنا أن الفئة ٦٠ – ٧٠ معناها ٦٠ وأقل من ٧٠ أي أن الطالب الحاصل على ٧٠ لا ينتمي لهذه الفئة ولكن ينتمي للفئة التالية ويستثنى من هذا التعريف الحد الأعلى للقيم إذا كان هو نفس الحد الأعلى للفئة الأخيرة فالفئة الأخيرة ٩٠ – ١٠٠ تشتمل على الحاصلين على مائة درجة في الاختبار.

بعد تحديد الفئات نبدأ بوضع كل قيمة في الفئة المناسبة لها عن طريق وضع علامات ### / كما أوضحنا في تبويب البيانات يدوياً مع ملاحظة تعريف الشرطة (-) كما أوردنا وقد ننتهي إلى جداول مثل الجدول التالي بعد إزالة العمود الأوسط الخاص بالعلامات لنحصل على التوزيع التكراري لدرجات مائة طالب في الصورة التالية :

الفئات	التكرار
٠ – ٦٠	١٥
٦٠ – ٧٠	٣٢
٧٠ – ٨٠	٢٨
٨٠ – ٩٠	١٦
٩٠ – ١٠٠	٩
المجموع	١٠٠

الجداول التكرارية المتجمعة

في كثير من الأحيان نحتاج إلى معرفة المفردات أو التكرارات التي أقل من قيمة معينة أو أكثر من قيمة معينة وللحصول على هذه التكرارات نستطيع تجميع التكرارات على التوالي أو تطريحها على التوالي للحصول على ما نسميه الجدول التكراري المتجمع الصاعد والجدول التكراري المتجمع الهابط.

والعملية ببساطة أن الجدول التكراري المتجمع الصاعد يبدأ دائماً بالتكرار صفر وينتهي بالمجموع عن طريق الجمع التتابعي والجدول التكراري المتجمع الهابط يبدأ بالمجموع وينتهي بالصفر عن طريق الطرح التتابعي، وذلك كما سيتضح من عمل الجدول المتجمع الصاعد والهابط للتوزيع التكراري للمثال السابق الخاص بدرجات المائة طالب.

الجدول الأصلي			الجدول التكراري المتجمع الصاعد			الجدول التكراري المتجمع الهابط	
فئات	تكرار		حدود عليا	تكرار متجمع صاعد		حدود دنيا	تكرار متجمع هابط
٠ ـ	١٥		-صفر	صفر		صفر فأكثر	١٠٠
٦٠ ـ	٣٢		-٦٠	١٥		٦٠ فأكثر	٨٥
٧٠ ـ	٢٨		-٧٠	٤٧		٧٠ فأكثر	٥٣
٨٠ ـ	١٦		-٨٠	٧٥		٨٠ فأكثر	٢٥
٩٠-١٠٠	٩		-٩٠	٩١		٩٠ فأكثر	٩
			-١٠٠	١٠٠		١٠٠ فأكثر	صفر
المجموع	١٠٠						

يلاحظ على هذه الجداول ما يلي :

١) الجدول الأصلي اختصاراً تكتب الفئات بدون الحد الأعلى حيث أنه الحدة الأدنى للفئة التالية ولعدم اللبس لأن الحد الأعلى يدخل ضمن الفئة التالية وذلك تبعاً لتعريف الشرطة (-) بأقل من الفئة الأولى تبدأ بصفر وتنتهي بأقل من ٦٠ سواء كان أقل من ٦٠ هو ٥٩ أو ٥٩,٥ أو حتى ٥٩,٥ كل ذلك يدخل في الفئة الأولى واعتباراً من ٦٠ يدخل في الفئة الثانية، وكذلك الفئة الثانية تبدأ بـ ٦٠ وتنتهي بأقل من ٧٠ سواء كان الرقم أقل من ٧٠ هو ٦٩ أو ٦٩,٥ أو ٦٩,٩ واعتباراً من ٧٠ يدخل ضمن الفئة الثالثة ويكون هذا التوزيع بهذه الكيفية صالح للقيم المنفصلة والقيم المتصلة على حد سواء.

٢) الرقم ٢٨ في الجدول الأصلي يعني وجود ٢٨ طالب درجاتهم في الاختبار النهائي تتراوح بين ٧٠ وأقل من ٨٠ أي أنه عدد الحاصلين على تقدير جيد بينما الرقم ٧٥ مثلاً في الجدول التكراري المتجمع الصاعد يعني وجود ٧٥ طالب حصلوا على درجة أقل من ٨٠ أي مجموع الطلاب الحاصلين على تقديرات جيد ومقبول وراسبين، في حين أن رقم ٥٣ في الجدول المتجمع الهابط يوضح عدد الطلاب الحاصلين على درجة ٧٠ فأكثر أي مجموع الطلاب الحاصلين على تقدير جيد وجيد جداً وممتاز، وهكذا يجب معرفة دلالة كل رقم سواء كان في الجدول التكراري الأصلي أو في الجدول المتجمع الصاعد أو الهابط.

٣) أحياناً يكون التوزيع التكراري الأصلي مفتوح من أعلى أو مفتوح من أسفل بمعنى ليس هناك حد أدنى للفئة الأولى وليس هناك حد أعلى للفئة الأخيرة ومع ذلك يمكن عمل الجدولين الصاعد والهابط بدون مشكلات بكتابة في الجدول الصاعد ـ الحد الأدنى للفئات (صفر)، ـ الحد الأعلى للفئات (المجموع) وفي الجدول المتجمع الهابط نكتب، الحد الأدنى فأكثر (المجموع) ـ الحد الأعلى فأكثر (صفر).

ثانياً : العرض البياني :

عادة ما تبدو الجداول والأرقام للعامة جافة وغير مشوقة، لذا يفضل دائماً لجذب الانتباه وشد النظر تجاه البيانات التعبير عنها في صورة ورسوم وخطوط بيانية وذلك طبعاً عن طريق الرسم. وبداية نحدد أن العرض البياني كله مسألة ذوق وفن أكثر منها علم وذلك لأن ما يمكن أن نستخلصه من الرسوم البيانية يمكن معرفته من الجداول لكنها وسيلة لجذب الانتباه .

الأساس الرياضي للتمثيل البياني :

يستعمل في التمثيل البيان محوران متعامدان المحور الأفقي ويطلــق عليــه المحور السيني والمحور الرأسي ويطلق عليه المحور الصادي. ويطلق علــى نقطة تقاطعهما نقطة الأصل وتكون دائماً مساوية للصفر وعلى المحور الأفقي تكون القيمة الموجبة يميناً والقيمة السالبة يساراً وعلى المحور الرأسي تكــون القيمة الموجبة فوق نقطة الأصل والقيمة السالبة تحت نقطة الأصل.

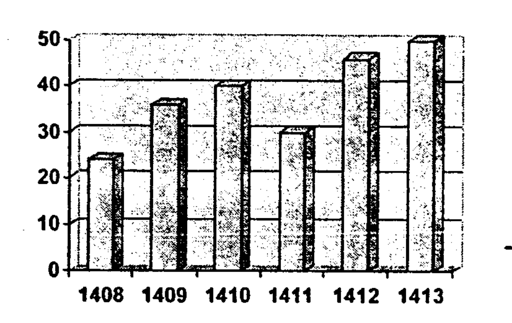
لوحة الأعمدة Bar Chart

في لوحة الأعمدة دائماً يكون المحور الأفقــي لأوجــه الظــاهرة المختلفــة والمحور الرأسي للتكرارات أو لقيم الظاهرة وتســتخدم كثيــراً فــي عــرض السلاسل الزمنية حيث يكون المحور الأفقي للسنوات والرأسي لقيم الظاهرة عبر السنوات المختلفة.

مثال (١) :

البيانات التالية تمثل تطور إنتاج السعودية من البترول بالمليون برميل.

السنة	١٤٠٨	١٤٠٩	١٤١٠	١٤١١	١٤١٢	١٤١٣
الإنتاج	٢٤	٣٦	٤٠	٣٠	٤٦	٥٠

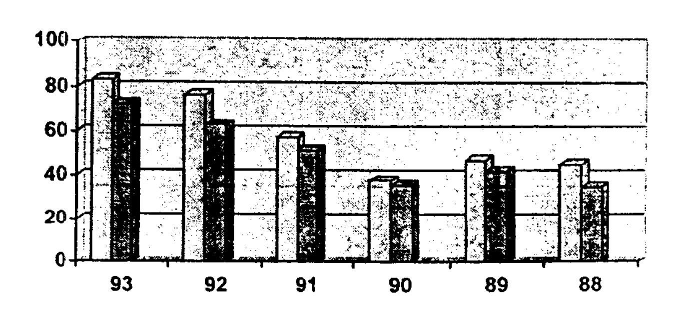


تطور إنتاج السعودية من البترول

مثال (٢) :

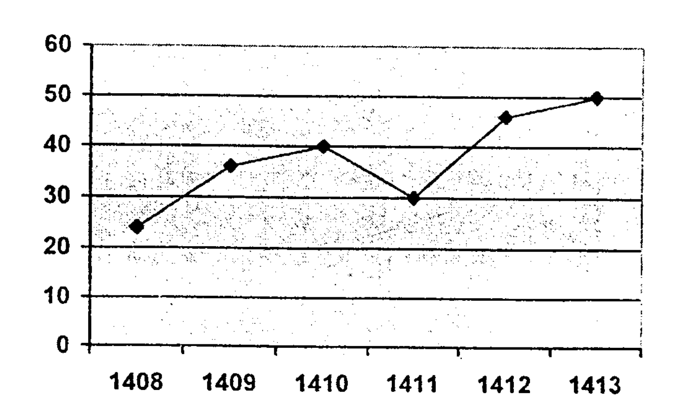
الجدول التالي يبين قيمة الصادرات وقيمة الواردات لدولة ما في الأعوام ١٩٨٦ – ١٩٩٣ بالمليون دولار :

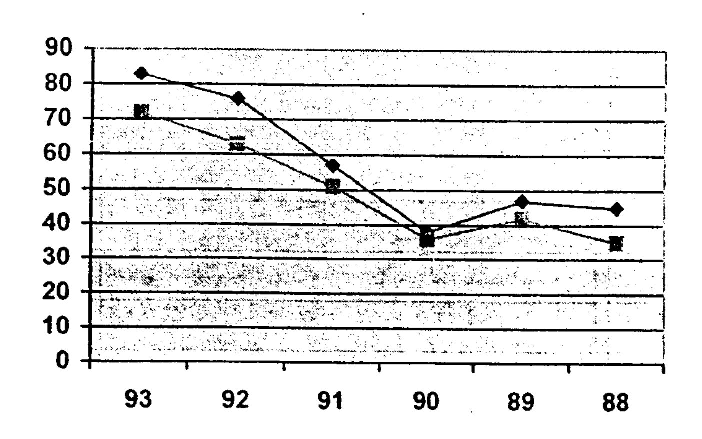
السنة	٨٨	٨٩	٩٠	٩١	٩٢	٩٣
الصادرات	٤٥	٤٧	٣٨	٥٧	٧٦	٨٣
الواردات	٣٥	٤٢	٣٦	٥١	٦٣	٧٢



الخط البياني

يمكن التعبير عن الجدولين السابقين برسم خط بياني كالتالي





يلاحظ على لوحة الأعمدة والخط البياني أن القارئ بمجرد النظر يســتطيع وبسرعة معرفة أن هناك هبوطاً في الإنتاج في سنة ١٤١١ وكذلك هبوط فــي الصادرات والواردات في عام ١٩٩٠م ومن هنا يبدأ التحليل لمعرفة الأسباب.

عرض البيانات بطريقة الدائرة Pie Chart :

كثيراً ما نرى دوائر مقسمة إلى أجزاء تستخدم لعرض البيانات وهــو مــا يطلق عليه عرض البيانات بطريق الدائرة Pie Chart وتستعمل الــدائرة فــي عرض ظاهرة واحدة مقسمة إلى أجزا أي دائماً يكون المجموع له معنى، مثال ذلك إنتاج دول الخليج من البترول مقسم إلى أجزاء لكل دولة جــزء، ونرســم الدائرة بعد تحديد نصيب كل جزء من الــ ٣٦٠ درجة هندســية التــي هــي مجموع درجات الدائرة وذلك بقسمة قيمة إنتاج كل دولة ÷ المجمــوع الكلــي والناتج يضرب في ٣٦٠ كما سيتضح في المثال التالي :

مثال (٣) :

الآتي بيان بإنتاج دول الخليج العربية من البترول فــي إحــدى الســنوات بالمليون برميل، أعرض هذه البيانات بطريقة الدائرة :

الدولة	السعودية	قطر	البحرين	الإمارات	الكويت	عمان	المجموع
الإنتاج	٤٢	٢٨	١٤	٥٤	٦٥	٣٢	٢٣٥

الحل :

درجات السعودية = (٤٢ ÷ ٢٣٥) × ٣٦٠ = ٦٤°

درجات قطر = (٢٨ ÷ ٢٣٥) × ٣٦٠ = ٤٣°

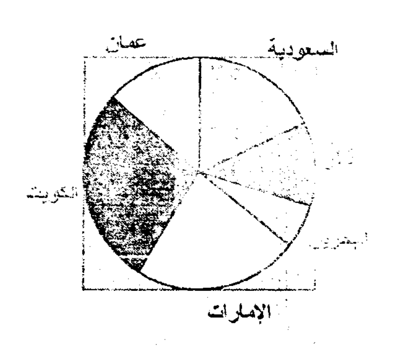
درجات البحرين = (١٤ ÷ ٢٣٥) × ٣٦٠ = ٢١°

درجات الأمارات = (٥٤ ÷ ٢٣٥) × ٣٦٠ = ٨٣°

درجات الكويت = (٦٥ ÷ ٢٣٥) × ٣٦٠ = ١٠٠°

درجات عمان = (٣٢ ÷ ٢٣٥) × ٣٦٠ = ٤٩°

المجموع = ٣٦٠°



العرض البياني للتوزيعات التكرارية :

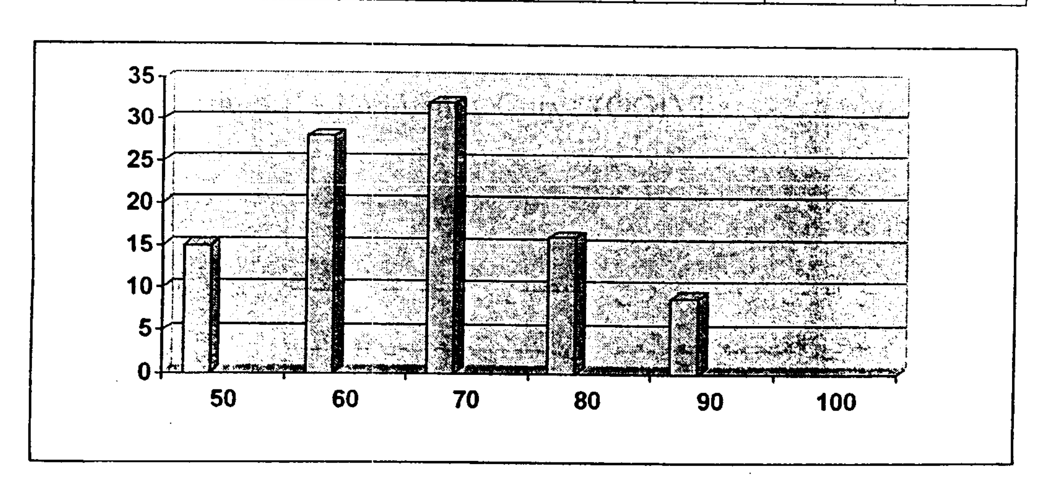
تعطينا الجداول التكرارية صورة عامة عن توزيع القيم ولكن الرسوم البيانية لهذه الجداول تعطينا صورة أسرع وأوضح خاصة لأغراض المقارنة ويستعمل في التمثل بالرسم طرق عديدة هي :

١. المدرج التكراري Frequency Histogram
٢. المضلع التكراري Frequency Polygon
٣. المنحنى التكراري Frequency Curve
٤. المنحنى التجميعي Cumulative Frequency Curve (OGIVE)

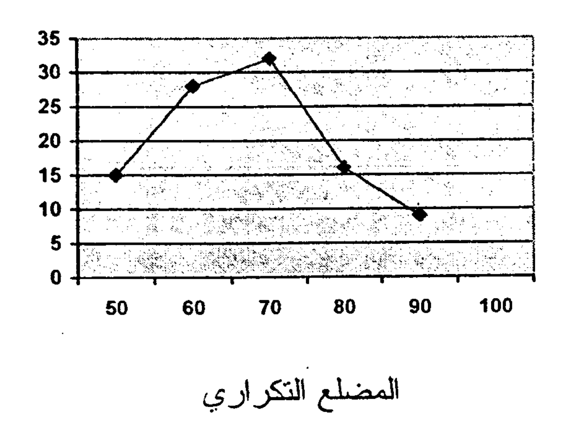
مثال (٤) :

لبيان توزيع التالي ارسم المدرج والمضلع والمنحنى التكراري وأيضاً المنحنى التكراري المتجمع مع العلم بأن البيانات تمثل توزيع أوزان مائة شخص.

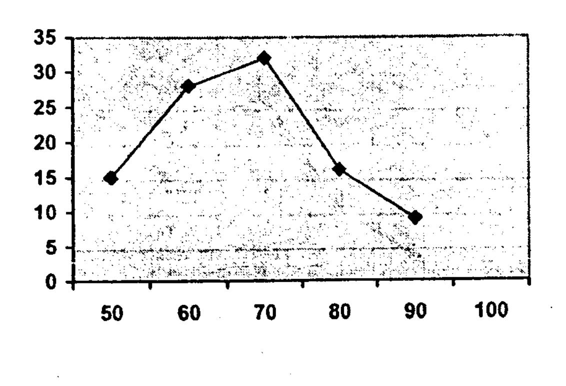
فئات	٥٠-	٦٠-	٧٠-	٨٠-	٩٠-١٠٠	مجموع
تكرار	١٥	٢٨	٣٢	١٦	٩	١٠٠



المدرج التكراري



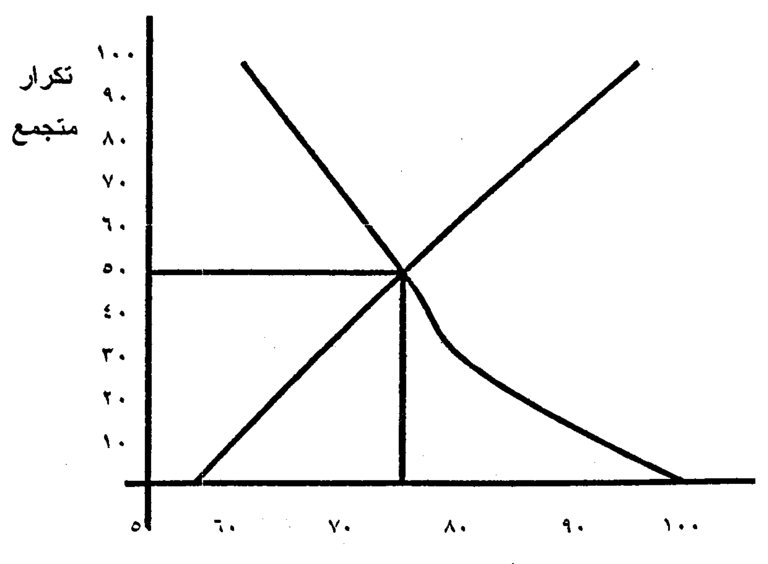
المضلع التكراري



المنحنى التكراري

لرسم المنحنى التكراري المتجمع (OGIVE) يلزمنا عمل توزيع تكـــراري متجمع صاعد كما سبق وأوردنا كالتالي :

حدود عليا	تكرار متجمع صاعد	حدود دنيا	تكرار متجمع هابط
٥٠-	صفر	٥٠ فأكثر	١٠٠
٦٠-	١٥	٦٠ فأكثر	٨٥
٧٠-	٤٣	٧٠ فأكثر	٥٧
٨٠-	٧٥	٨٠ فأكثر	٢٥
٩٠-	٩١	٩٠ فأكثر	٩
١٠٠-	١٠٠	١٠٠ فأكثر	صفر



المنحنى التكراري المتجمع الصاعد ــ الهابط

الوسيط = ٧٢

الفصل الخامس
مقاييس النزعة المركزية
Measures of Central Tendency

يقصد بالنزعة المركزية للبيانات الإحصائية ميل هذه البيانات إلى التراكم أو التجمع حول قيمة متوسطة معينة للظاهرة موضع الدراسة وتسمى هذه القيمة التي تتجمع حولها البيانات بالقيمة المتوسطة Average Value. هذا ويجب أن يتوفر في المتوسط المقدر الشروط التالية :

١) يجب ألا يتأثر بالقيم الشاذة أو المتطرفة.
٢) يجب أن يأخذ جميع مفردات الظاهرة في الاعتبار عند حسابه، حيث أن إهمال بعض المفردات يجعل المتوسط غير ممثل للظاهرة تمثيلاً صحيحاً.
٣) يجب ألا يتأثر قدر الإمكان بتغير العينات المسحوبة من نفس المجتمع. ومما هو جديد بالذكر فإنه من المتوقع أن تختلف قيم متوسطات العينات المسحوبة مهما كانت درجة الدقة في اختيارها غير أنه يتعين اختيار ذلك المتوسط الذي يجعل مثل هذه الاختلافات أقل ما يمكن.
٤) سهولة فهمه ووضوح معناه ومدلوله.
٥) سهولة وسرعة حسابه وخضوعه للعمليات الجبرية المعتادة.

وهناك العديد من مقاييس النزعة المركزية والتي من أهمها:

أولاً : الوسط الحسابي (المتوسط الحسابي) Arithmetic Mean

ثانياً : الوسيط Median

ثالثاً : المنوال Mode

رابعاً : الوسط الهندسي Geometric Mean

وسوف نتناول فيما يلي كيفية تطبيق تلك المقاييس سواء بالنسبة للبيانات غير المبوبة أو تلك المبوبة في جداول توزيع تكراري.

الوسط الحسابي Arithmetic Mean :

الوسط الحسابي لمجموعة من القيم هو عبارة عن حاصل جمع هذه القيم مقسوماً على عدد المفردات :

مثال (١) :

لو كان لدينا الأجور اليومية لسبع مشتغلين كما يلي :

٨، ١٠، ١٧، ١٢، ١٤، ١٦، ١٤ جنيه

فيكون الوسط الحسابي لأجور هؤلاء المشتغلين السبعة هو:

$$= \frac{\text{مجموع الأجور}}{\text{عدد المشتغلين}} = \frac{٩١}{٧} = ١٣ \text{ جنيه}$$

ويلاحظ من هذا المثال أن أجور المشتغلين هي ظاهرة متغيرة لأن الأجر يتغير من مشتغل إلى أخر. فلو رمزنا للأجور (قيم المتغير) بالرمز س فإنه يكون لدينا :

$س_١ = ٨$ $س_٢ = ١٠$ $س_٣ = ١٧$ $س_٤ = ١٢$ $س_٥ = ١٤$

$س_٦ = ١٦$ $س_٧ = ١٤$

ويكون الوسط الحسابي هو :

$$= \frac{س_1 + س_2 + س_3 + س_4 + س_5 + س_6 + س_7}{7}$$

$$= \frac{(مجـ\ س)}{7}$$

ويمكن تعميم المعادلة السابقة لحساب الوسط الحسابي لأي ظاهرة يمثلها متغير ما (س) وعدد مفرداتها (ن) وذلك بفرض أن الوسط الحسابي هو $\bar{س}$، فإن :

$$\bar{س} = \frac{س_1 + س_2 + ... + س_ن}{ن}$$

$$\bar{س} = \frac{(مجـ\ س)}{ن} \quad .... (1)$$

ويمكن أن نستنتج من المعادلة (١) أن :

$$مجـ\ س = \bar{س} \times ن \quad .... (2)$$

وللتعرف على طريقة إيجاد الوسط الحسابي في التوزيعات التكرارية بصفة عامة ينبغي الإلمام بخصائص الوسط الحسابي. وهو ما سوف نتناوله فيما يلي:

خصائص الوسط الحسابي

لما كان الوسط الحسابي يخضع للبرهنة الرياضـــية فقــد أمكــن معرفـــة خصائصه والتي تتضمن :

(١) المجموع الجبري لإنحرافات القيم عن وسطها الحسابي يساوي صفراً.

ويمكن الوصول إلى ذلك لو أخذنا في الاعتبار المثال الســـابق والخـــاص بأجور ٧ مشتغلين وهي : ٨، ١٠، ١٧، ١٢، ١٤، ١٦، ١٤ جنيه حيث نجــد أن:

س	$\bar{س}$	(س – $\bar{س}$) الانحرافات
٨		٥–
١٠		٣–
١٧		٤
١٢	١٣	١–
١٤		١
١٦		٣
١٤		١
٩١		٠

كما يمكن البرهنة رياضياً على هذه الخاصية للوسط الحسابي على الوجـــه التالي :

$\because$ مجـ (س – $\bar{س}$) = مجـ س – مجـ $\bar{س}$

$\because$ مجــ س – مجــ $\overline{س}$ = مجــ $\overline{س}$ (ن . $\overline{س}$)

$\therefore$ مجــ س (ن. $\overline{س}$) = مجــ س – (ن . $\overline{س}$)

(٢)إذا أضفنا إلى قيم مفردات المتغير (س) أو طرحنا منها كمية ثابتة (أ).
مثلاً فإنه يكون لدينا قيم مفردات جديدة لمتغير جديد (ص). إن الوسط الحسابي لقيم مفردات المتغير (س) هو عبارة عن الوسط الحسابي لقيم مفردات المتغير (ص) مطروحاً منه أو مضافاً إليه الكمية الثانية (أ).
وهذه الخاصية يمكن الاستفادة منها في حالة ما إذا كانت قيم مفردات المتغير (س) قيم كبيرة ويخشى الخطأ في جمعها لإيجاد الوسط الحسابي.

مثال (٢) :

إذا كان المطلوب إيجاد الوسط الحسابي لمجموعة القيم الآتية :

٩٦٨٨٧٠	٩٦٨٩٥٠
٩٦٨٨٩٠	٩٦٨٩٧٠
٩٦٨٩٢٠	

لإيجاد الوسط الحسابي بالنسبة لهذه القيم الكبيرة ذات المدى الضئيل يمكن اختيار قيمة منها ككمية ثابتة (أ) وطرحها من القيم الموجودة بحيث نحصل على قيم متغير جديد (ص) أي أن:

$س_{١} - أ = ص_{١}$

$س_{٢} - أ = ص_{٢}$

$س_{٣} - أ = ص_{٣}$

وتكون قيمة $\overline{س}$ = $\overline{ص}$ + أ

وبفرض أن أ = ٩٦٨٨٧٠ فإنه يكون لدينا :

٩٦٨٨٧٠ – ٩٦٨٨٧٠ = ٠

٩٦٨٨٩٠ – ٩٦٨٨٧٠ = ٢٠

٩٦٨٩٢٠ – ٩٦٨٨٧٠ = ٥٠

٩٦٨٩٥٠ – ٩٦٨٨٧٠ = ٨٠

٩٦٨٩٧٠ – ٩٦٨٨٧٠ = ١٠٠

وحيث أن $\overline{ص}$ = $\frac{٢٥٠}{٥}$ = ٥٠

∴ $\overline{س}$ = ٥٠ + ٩٦٨٨٧٠ = ٩٦٨٩٢٠

ويلاحظ في المثال السابق أننا أضفنا الكمية الثابتة على قيمة الوسط الحسابي للمتغير الجديد وذلك لأننا قمنا بطرح هذه الكمية الثابتة من جميع قيم المتغير (س) ولذلك فإننا نجد أنه في الحالة العكسية وهي التي نقوم فيها بإضافة كمية ثابتة إلى قيم المتغير (س)، أن قيمة الوسط الحسابي للمتغير (س) : $\overline{س}$ = قيمة الوسط الحسابي للمتغير الجديد ص مطروحاً منها الكمية الثابتة أي أن س = ص – ١

وعلى ذلك فإننا نجد في حالتي الإضافة أو الطرح من قيم المتغير (س) أن الوسط الحسابي س = ص ± أ

ويمكن البرهنة على ذلك كما يأتي :

$ص_١ = س_١ \pm أ$

$ص_٢ = س_٢ \pm أ$

$ص_٣ = س_٣ \pm أ$

....

....

....

$ص_ن = س_ن + أ$

وتكون $ص = س \pm أ$

، $مجـ ص = مجـ (س \pm أ)$

أ، $مجـ ص = مجـ س \pm ن أ$

وبقسمة طرفي المعادلة على ن نجد أن :

$$\frac{مجـ ص}{ن} = \frac{مجـ س}{ن} \pm أ$$

$\therefore \overline{ص} = \overline{س} \pm أ$

$\overline{س} = \overline{ص} \pm أ$ .... **(٣)**

(٣)إذا ضربنا قيم المتغير (س) في كمية ثابتة (ب)، فإنه يكون لدينا قيم لمفردات جديدة لمتغير جديد (ص). ويكون الوسط الحسابي لقيم المتغير (س) هو عبارة عن الوسط الحسابي لقيم المتغير (ص) مقسوماً على الكمية الثابتة (ب).

والمثال التالي يوضح هذه الخاصية من خصائص الوسط الحسابي والتي يمكن الاستفادة منها في حالة القيم ذات الكسور.

مثال (٣) :

أوجد الوسط الحسابي للقيم :

$$\frac{٣٨}{٧} \text{ ، } \frac{٦}{٧} \text{ ، } \frac{٦١}{٢١} \text{ ، } \frac{٣٢}{٢١} \text{ ، } \frac{٢٥}{٢١}$$

لإيجاد الوسط الحسابي لهذه القيم ذات الكسور فإنه يمكن اختيار كمية ثابتة (ب) وضربها في القيم الموجودة بحيث تحصل على قيم متغيرة جديدة (ص) ويشترط أن تكون هذه القيم هي أعداد صحيحة، ففي مثالنا هذا يمكن اختيار (ب) = ٢١ وبضربها في قيم المتغير (س) نجد أن :

س	ب	ص
$\frac{٣٨}{٧}$	× ٢١ =	١١٤
$\frac{٦}{٧}$	× ٢١ =	١٨
$\frac{٦١}{٢١}$	× ٢١ =	٦١
$\frac{٣٢}{٢١}$	× ٢١ =	٣٢
$\frac{٢٥}{٢١}$	× ٢١ =	٢٥

$$\frac{\text{مجـ ص}}{\text{ن}} \quad \frac{٢٥٠}{٥}$$

$\therefore \bar{ص} = \quad = \quad = ٥٠$

$\bar{س} = \frac{ص}{ب} = \frac{٥٠}{٢١} = ٢,٣٨$

ويمكن البرهنة على ذلك إذا فرضنا أن قيم المتغير (س) عددها (ن) وأن :

$ص_{١} = س_{١} \times ب$

$ص_{٢} = س_{٢} \times ب$

$ص_{٣} = س_{٣} \times ب$

....

....

$ص_{ن} = س_{ن} \times ب$

وتكون $ص = س \times ب$

$\therefore$ مجـ ص = مجـ س × ب

وبقسمة طرفي المعادلة على ، نجد أن :

$$\frac{مجـ ص}{ن} = \frac{مجـ س}{ن} \times ب$$

$\therefore \bar{ص} = \bar{س} \times ب$

$$\bar{س} = \frac{\bar{ص}}{ب} \quad ....(٤)$$

(٤)إذا قسمنا قيم المتغير (س) على كمية ثابتة (جـ) فإنه يكون لدينا قيم لمفردات جديدة لمتغير جديد (ص). ويكون الوسط الحسابي لقيم المتغير (س) هو عبارة عن الوسط الحسابي لقيم المتغير (ص) مضروباً في الكمية الثابتة (جـ).

والمثال الآتي يوضح هذه الخاصية :

مثال (٤) :

أوجد الوسط الحسابي للقيم :

١٢٥٠٠ ، ٧٥٠٠ ، ١٥٠٠٠ ، ١٠٥٠٠ ، ١٣٠٠٠

بالنسبة لمثل هذه القيم الكبيرة والتي يكون المدى فيها كبيراً أيضاً فلإيجاد الوسط الحسابي لها فإنه يمكن اختيار كمية ثابتة (جـ) وقسمة جميع القيم عليها بحيث تحصل على قيمة متغير جديد (ص) ويلاحظ أن تقبل جميع القيم الموجودة القسمة على الكمية الثابتة.

وفي مثالنا يمكن اعتبار الكمية الثابتة = ٥٠٠ وبقسمة القيم عليها ينتج أن :

س		جـ		ص
١٢٥٠٠	÷	٥٠٠	=	٢٥
٧٥٠٠	÷	٥٠٠	=	١٥
١٥٠٠٠	÷	٥٠٠	=	٣٠
١٠٥٠٠	÷	٥٠٠	=	٢١
١٣٠٠٠	÷	٥٠٠	=	٢٦
				١١٧

$$\frac{\text{مجـ ص}}{\text{ن}} \quad \frac{١١٧}{٥}$$

∴ $\bar{ص}$ = = = ٢٣,٤

وتكون $\bar{س}$ = $\bar{ص}$ × جـ = ٢٣,٤ × ٥٠٠ = ١١٧٠٠

ويمكن البرهنة على ما سبق لو افترضنا أن قيم المتغير س عددها ن وأن :

$ص_١$ = $س_١$ ÷ جـ

$ص_٢$ = $س_٢$ ÷ جـ

$ص_٣$ = $س_٣$ ÷ جـ

....

....

....

$ص_ن$ = $س_ن$ ÷ جـ

وأن ص = $\frac{س}{جـ}$

، مجـ ص = مجـ س × $\frac{١}{جـ}$

وبقسمة طرفي المعادلة على ن ينتج أن :

$\bar{ص}$ = $\frac{\bar{س}}{ج}$

∴ $\bar{س}$ = ص × جـ .... (٥)

(٥) إذا طرحنا من قيم المتغير (س) كمية ثابتة (أ) وقسمن الناتج على كمية ثابتة (جـ) فإنه يكون لدينا قيم لمفردات جديدة لمتغير جديد (ص).

ويكون الوسط الحسابي لقيم المتغير (س) هو عبارة عن الوسط الحسابي لقيم المتغير (ص) مضروباً في الكمية الثابتة (جـ) ومضافاً إلى الناتج الكمية الثابتة (أ).

والخاصية السابقة للوسط الحسابي توصلنا إليها باستخدام الخاصيتين الثانية والرابعة، ويستفاد منها عند تبسيط الأعداد ذات القيم الكبيرة والتي يكون المدى فيها كبيراً كما يتضح من المثال الآتي :

مثال (٥) :

المطلوب إيجاد الوسط الحسابي للقيم الآتية :

٩٣٩٧٢٥ ، ٩٤٠٢٣٠

٩٤٠٧٢٥ ، ٩٤١٢٤٠

٩٤١٧٤٥

لإيجاد الوسط الحسابي في مثل هذه الحالة نقوم بطرح كمية ثابتة (أ) ولتكن ٩٣٩٧٢٥ وتسمى هذه القيمة بالوسط الفرضي. ثم نقسم بواقي الطرح على كمية ثابتة (جـ) ولاختيار هذه الكمية الثابتة يشترط أن تقبل جميع القيم القسمة عليها. ونجد أن القاسم المشترك في هذا المثال هو ٥٠٠ وعلى ذلك يكون لدينا:

س	س - ١	ص = (س - ١) / جـ
٩٣٩٧٢٥	٠	٠
٩٤٠٢٣٠	٥٠٠	١
٩٤٠٧٢٥	١٠٠٠	٢
٩٤١٢٤٠	١٥٠٠	٣
٩٤١٧٤٥	٢٠٠٠	٤
		١٠

$$\therefore \bar{ص} = \frac{مجـ ص}{ن} = \frac{١٠}{٥} = ٢$$

$$\because \bar{س} = (\bar{ص} \times جـ) + أ$$

$$\bar{س} = (٢ \times ٥٠٠) + ٩٣٩٧٢٥$$

$$\bar{س} = ٩٤٠٧٢٥$$

ويمكن البرهنة على ما تقدم لو افترضنا أن قيم المتغير س عددها (ن) وأن:

$$ص_{١} = \frac{س_{١} - ا}{جـ}$$

$$ص_{٢} = \frac{س_{٢} - ا}{جـ}$$

$$ص_{ن} = \frac{س_{ن} - ١}{جـ}$$

وأن $ص = \frac{س - ١}{جـ}$

جـ ص = س – أ

س = ج ص + أ

∴ مجـ س = جـ مجـ ص + ن أ

وبقسمة طرفي المعادلة على ن ينتج أن :

$\bar{س}$ = جـ $\bar{ص}$ + أ

وبعد أن استعرضنا فيما سبق أهم خصائص الوسط الحسابي، سوف نتناول فيمل يلي طريقة إيجاد الوسط الحسابي في الجداول التكرارية بصفة عامة وسواء أكان ذلك وفقاً للطريقة العادية أو باستخدام الخواص السابقة منتهين إلى الطريقة المختصرة لإيجاد الوسط الحسابي.

(٢) الوسط الحسابي في حالة الجداول التكرارية :

(I) إيجاد الوسط الحسابي دون استخدام خواصه :

(أ) الوسط الحسابي للجداول التكرارية البسيطة :

لإيجاد الوسط الحسابي إذا كان التوزيع ممثلاً في جدول تكراري فلابد وأن يؤخذ في الاعتبار قيم التكرارات. والمثال التالي يوضح طريقة إيجاد الوسط الحسابي في التوزيع التكراري البسيط.

مثال (٦) :

فيما يلي توزيع تكراري بسيط لعدد ٤١٥ تاجراً وفقاً لمبيعاتهم اليومية :

المبيعات (س)	عدد التجار (ك)
٥	٨
١٠	٢٦
١٥	٩٥
٢٠	١٦٠
٢٥	٨٤
٣٠	٣١
٣٥	١١
المجموع	٤١٥

والمطلوب إيجاد الوسط الحسابي لهذا التوزيع.

يلاحظ في هذا المثال أن كل قيمة للمبيعات حققها عدد من التجار فمثلاً القيمة (٥) مكررة ٨ مرات والقيمة ١٠ مكررة ٣٦ مرة وهكذا، فلإيجاد الوسط الحسابي لابد من الحصول على مجموع القيم المختلفة. وللحصول على هذا المجموع لابد من إيجاد مجموع كل قيمة على حدة، فمجموع القيمة ٥ هو ٥ × ٨ ومجموع القيمة ١٠ هو ١٠ × ٢٦، وهكذا حتى نهاية القيم، ويكون مجموع حواصل الضرب عبارة عن مجموع القيم المختلفة في تكراراتها . وهذا يعني أنه أصبح للتكرارات وزن في حساب الوسط الحسابي.

وبتطبيق ما سبق نصل إلى الجدول التالي :

س	ك	س.ك
٥	٨	٤٠
١٠	٢٦	٢٦٠
١٥	٩٥	١٤٢٥
٢٠	١٦٠	٣٢٠٠
٢٥	٨٤	٢١٠٠
٣٠	٣١	٩٣٠
٣٥	١١	٣٨٥
المجموع	٤١٥	٨٣٤٠

وحيث أن $\bar{س} = \frac{مجـ س}{ن}$ من القانون (١)

فإن الوسط الحسابي في الجداول التكرارية البسيطة هو :

$$\bar{س} = \frac{مجـ (س.ك)}{ن}$$ .... (١ مكرر)

$$\therefore \bar{س} = \frac{٨٣٤٠}{٤١٥} = ٢٠,٠١$$

(ب) الوسط الحسابي للجداول التكرارية ذات الفئات المتساوية :

في حالة الجداول التكرارية ذات الفئات المتساوية نجد أن هناك فئات متساوية، وكل فئة يكون أمامها عدد من التكرارات، وكما سبق أن أوضحنا فإنه يصعب معرفة القيمة الأصلية لأي مفردة من الجدول ولا نعرف عنها إلا أنها واحدة من فئة معينة بين حدود معينة. لهذا فلإيجاد الوسط الحسابي نعتبر أن قيمة أي مفردة من المفردات التي تقع في فئة واحدة هي مركز هذه الفئة، ونعتبر أن مركز الفئة يتكرر بعدد المفردات التي تمثل الفئة، ويسهل بعد ذلك إيجاد مجموع القيم على الوجه الذي سبق أن أوضحناه بالنسبة للجداول التكرارية البسيطة وفيما يلي مثالاً يوضح ذلك.

مثال (٧) :

فيما يلي توزيع تكراري لعدد ٣٠٠ مشتغلاً بالشركة العربية للصناعات الخفيفة وفقاً لفئات أجورهم اليومية :

الفئات	التكرارات
٥ – ١٠	٢٢
١٠ – ١٥	٤٣
١٥ – ٢٠	٥٠
٢٠ – ٢٥	١٢٠
٢٥ – ٣٠	٣٥
٣٠ – ٣٥	٢٠
٣٥ – ٤٠	١٠
المجموع	٣٠٠

والمطلوب إيجاد الوسط الحسابي لهذا التوزيع.

يلاحظ أن هذا الجدول يختلف عن الجدول التكراري البسيط حيث توجد في هذا الجدول فئات. فمثلاً في الفئة من ١٠ – ١٥ يوجد ٤٣ مشتغلاً أجر كل منهم غير محدود.

لهذا فإننا نعتبر أن أجر كل عامل في أي فئة من الفئات السابقة هو عبارة عن مركز الفئة. فمثلاً نعتبر أن أجر كل مشتغل من الـ ٢٢ مشتغلاً في الفئة الأولى هو ٧,٥ جنيه، ٧,٥ج تمثل مركز الفئة $\frac{١٠ + ٥}{٢}$ وأجر كل مشتغل من الـ ٤٣ مشتغلاً في الفئة الثانية هو ١٢,٥ ج، ولو رمزنا لمراكز الفئات بالرمز (س) والتكرارات بالرمز ك فإنه من السهل الحصول على مجموع القيم المختلفة لإيجاد الوسط الحسابي كما يتضح من الجدول التالي :

س	ك	س.ك
٧,٥	٢٢	١٦٥
١٢,٥	٤٣	٥٣٧,٥
١٧,٥	٥٠	٨٧٥
٢٢,٥	١٢٠	٢٧٠٠
٢٧,٥	٣٥	٩٦٢,٥
٣٢,٥	٢٠	٦٥٠
٣٧,٥	١٠	٣٧٥
المجموع	٣٠٠	٦٢٦٥

وباستخدام القانون (١ مكرر) نجد أن :

$$\bar{س} = \frac{مجـ (س.ك)}{ن} = \frac{٦٢٦٥}{٣٠٠} \qquad \bar{س} = ٢٠,٨٨٢ \text{ مليم جنيه}$$

ويلاحظ أخيراً أن القواعد السابقة تطبق فقط في حالة الجداول التكرارية المقفلة أما الجداول التكرارية المفتوحة أو التي تتضمن فئات غير محدد حدها الأدنى أو حدها الأعلى فإنه من المستحيل تحديد مركز الفئة اللازم لإيجاد الوسط الحسابي وبالتالي يستحيل التوصل إلى الوسط الحسابي.

(جـ) الوسط الحسابي للجداول التكرارية ذات الفئات غير المتساوية :

نطبق نفس القواعد السابق تطبيقها بالنسبة لإيجاد الوسط الحسابي في الجداول التكرارية ذات الفئات المتساوية. وهذه القواعد تطبق فقط في الجداول التكرارية المقفلة ذات الفئات غير المتساوية.

(II) إيجاد الوسط الحسابي باستخدام خواصه :

(أ) التوزيعات التكرارية البسيطة :

يمكن إيجاد الوسط الحسابي بتطبيق القوانين المبرهنة على خواصه والسابق ذكرها .

واستخدام هذه القوانين يؤدي إلى تبسيط العمليات الحسابية وخاصة إذا كانت الأرقام كبيرة، حيث أننا نقوم بطرح كمية ثابتة (أ) من جميع قيم المتغير (س) تسمى الوسط الفرضي، ويشترط في الوسط الفرضي أن يكون في وسط الجدول التكراري ومقابل أكبر تكرار إذا أمكن. ثم نقسم على القاسم المشترك إذا وجد في حالة ما إذا كانت قيم المتغير (س) تتزايد بأساس ثابت.

والمثال التالي يوضح طريقة حساب الوسط الحسابي في التوزيعات التكرارية البسيطة.

مثال (٨) :

فيما يلي توزيع عدد ٥٠٠ طالباً من طلبة السنة الثالثة بكلية التجارة وفقاً لدرجاتهم في مادة الإحصاء في امتحان مايو ١٩٩٣.

الدرجة	عدد الطلبة
١٠	١٠
٢٠	٢٠
٣٠	٢٥
٤٠	١١٥
٥٠	١٥٠
٦٠	١٠٠
٧٠	٥٠
٨٠	٢٠
٩٠	١٠
المجموع	٥٠٠

والمطلوب إيجاد الوسط الحسابي لهذا التوزيع

يمكن في هذا المثال اختيار الكمية الثابتة (أ) أو الوســـط الفرضـــي = ٥٠، والكمية الثابتة (جـ) أو القاسم المشترك = ١٠.

وحتى يمكن تطبيق القانون (٦) لابد من إعداد جدول كما يلي :

س	ك	س ــ أ	ص = (س ــ أ) / جـ	ك . ص
١٠	١٠	- ٤٠	- ٤	- ٤٠
٢٠	٢٠	- ٣٠	- ٣	- ٦٠
٣٠	٢٥	- ٢٠	- ٢	- ٥٠
٤٠	١١٥	- ١٠	- ١	- ١١٥
٥٠	١٥٠	٠	٠	٠
٦٠	١٠٠	١٠	١	١٠٠
٧٠	٥٠	٢٠	٢	١٠٠
٨٠	٣٠	٣٠	٣	٦٠
٩٠	١٠	٤٠	٤	٤٠
المجموع	٥٠٠			٣٥

$$\therefore \bar{ص} = \frac{مجـ(ك.ص)}{ن} = \frac{٣٥}{٥٠٠} = ٠,٠٧$$

$$\Theta \bar{س} = جـ ص + أ$$

$$\therefore \bar{س} = ١٠ \times ٠,٠٧ + ٥٠ = ٥٧$$

(ب) التوزيعات التكرارية ذات الفئات المتساوية :

ويلاحظ بالنسبة للتوزيعات التكرارية ذات الفئـــات المتســـاوية أن الوســـط الفرضي (الكمية الثابتة أ) هو مركز الفئة التي تقع في وسط التوزيع ويمثلهـــا أكبر تكرار إذا أمكن وأن القاسم المشترك (الكمية الثابتة جـــ) هو طول الفئـــة. وتتبع نفس الخطوات السابق بيانها بالنسبة للتوزيعات التكرارية البسيطة والمثال التالي يوضح ذلك.

مثال (٩) :

فيما يلي توزيع ٣٠٠ مصنعاً وفقاً لفئات رأس المال المستثمر بكـــل منهـــا بملايين الجنيهات :

رأس المال المستثمر بملايين الجنيهات	عدد المصانع
٥ – ١٠	٢٢
١٠ – ١٥	٤٣
١٥ – ٢٠	١٢٠
٢٠ – ٢٥	٥٠
٢٥ – ٣٠	٣٥
٣٠ – ٣٥	٢٠
٣٥ – ٤٠	١٠
المجموع	٣٠٠

في هذا المثال نجد أن الوسط الفرضي هو مركز الفئة التي أمامها أكبر تكرار وهو ١٧,٥ وأن القاسم المشترك هو طول الفئة أي ٥، ولإيجاد الوسط الحسابي يمكن إعداد جدول بالشكل الآتي :

س	ك	س ـ أ	ص = (س ـ أ) / جـ	ك . ص
٧,٥	٢٢	- ١٠	- ٢	- ٤٤
١٢,٥	٤٣	- ٥	- ١	- ٤٣
١٧,٥	١٢٠	٠	٠	٠
٢٢,٥	٥٠	٥	١	٥٠
٢٧,٥	٣٥	١٠	٢	٧٠
٣٢,٥	٢٠	١٥	٣	٦٠
٣٧,٥	١٠	٢٠	٤	٤٠
المجموع	٣٠٠			١٣٣

$$\therefore \bar{ص} = \frac{مجـ (ك.ص)}{ن} = \frac{١٣٣}{٥٠٠} = ٠,٤٤٣$$

$$\because \bar{س} = جـ \bar{ص} + أ$$

$$\therefore \bar{س} = ٥ \times ٠,٤٤٣ + ١٧,٥ = ١٩,٧١٥$$

وهذا يعني أن متوسط رأس المال المستثمر في هذه المصانع هو ١٩,٧١٥ جنيهاً ومن هذا المثال يمكن ملاحظة :

- أننا دائماً سنحصل على قيمة (ص) = صفر أمام الفئة ذات أكبر تكرار، ثم تأتي القيمة الموجبة ١، ٢، ٣ ... الخ للفئات التي تلي هذه الفئة في القيمة، وتأتي القيمة السالبة −١، −٢، −٣ ... الخ للفئات التي تسبق هذه الفئة في القيمة.

- أن الوسط الفرضي (أ) هو قيمة مركز الفئة التي اخترناها ووضعنها أمامها ص = صفر.

- أن القاسم المشترك (جـ) يمثله طول الفئة.

وعلى أساس هذه الملاحظات يمكن إيجاد الوسط الحسابي بطريقة مختصرة حيث يمكن الاستغناء عن الخانتين المخصصتين لمراكز الفئات وانحرافات القيم عن الوسط الفرضي. وعلى هذا فإنه لإعداد الجدول الخاص بإيجاد الوسط الحسابي يكتفي بأربع خانات للفئات والتكرارات وقيم المتغير الجديد (ص) وقيم (ك ص) والمثال التالي يوضح ذلك.

مثال (١٠) : فيما يلي توزيع ١٠٠ مشتغلاً وفقاً لدرجاتهم في اختبار القدرات نهايته العظمى ٢٠ درجة.

الفئات	التكرارات
٤ – ٦	٦
٦ – ٨	١٤
٨ – ١٠	٢٤

١٠ – ١٢	٣٢
١٢ – ١٤	١٦
١٤ – ١٦	٨
المجموع	١٠٠

وفي المثال يمكن وضع قيم المتغير (ص) مباشرة دون الالتجاء إلى مراكز الفئات أو تحديد قيم انحرافات المتغير (س) عن الوسط الفرضي.

وذلك باعتبار أن الفئة التي يمثلها أكبر عدد من التكرارات هي الفئة ١٠ – ١٢ ويمثلها ٣٢ تكراراً، فإنه يمكن وضع قيمة (ص) أمام هذه الفئة = صفراً. ثم نحدد قيم (ص) للفئات التي تلي هذه الفئة وهي ١، ٢ على التوالي، وكذلك قيم (ص) للفئات التي تسبق هذه الفئة وهي -١، -٢، -٣ على التوالي حيث يكون لدينا الجدول التالي :

س	ك	ص	ك . ص
٤ – ٦	٦	- ٣	- ١٨
٦ – ٨	١٤	- ٢	- ٢٨
٨ – ١٠	٢٤	- ١	- ٢٤
١٠ – ١٢	٣٢	٠	٠
١٢ – ١٤	١٦	١	١٦
١٤ – ١٦	٨	٢	١٦
المجموع	١٠٠		- ٣٨

∴ $\bar{ص} = \frac{مجـ (ك.ص)}{ن} = - ٠,٣٨$

وحيث أن أ = ١١ ، جـ = ٢

∴ $\bar{س}$ = (جـ $\bar{ص}$) + أ

س = (٢٠ × ٠,٣٨) + ١١

$\bar{س}$ = ١١ - ٠,٧٦ = ١٠,٢٤

وهذا يعني أن المتوسط الحسابي لدرجات هؤلاء المشتغلين في اختبار القدرات هو ١٠,٢٤.

٢) الوسيط Median :

هو قيمة المفردة التي تجعل عدد المفردات قبلها مساوياً لعدد المفردات بعدها شريطة أن ترتب المفردات ترتيباً تصاعدياً أو تنازلياً.

كيفية تحديد الوسيط من بيان غير مبوب :

مثال (١) :

إذا كان لدينا رأس مال (٩) شركات بالمليون جنيه في الصورة :

٦,٥ ، ٨ ، ٤ ، ٧ ، ٩,٢ ، ٣,٥ ، ٥ ، ٨,٧ ، ١٠

فإن وسيط رأس المال كأحد مقاييس المتوسطات يتطلب أولاً الترتيب تصاعدياً أو تنازلياً كالتالي :

٣,٥ ، ٤ ، ٥ ، ٦,٥ ، [٧] ، ٨ ، ٩,٥ ، ١٠

فإذا كانت عدد المفردات فردياً كما هو معطي فإن الوسيط هو قيمة المفردة التي تجعل عدد المفردات قبلها مساوياً لعدد المفردات بعدها وتحدد هذه المفردة بالموقع

$$\frac{ن + ١}{٢} = \frac{٩ + ١}{٢} = ٥$$

∴ المفردة الخامسة = ٧ هي قيمة الوسيط

فإذا كان عدد المفردات زوجياً، تصبح هناك مفردتـــان موقعهمـــا يتحـــدد بالقوانين $\frac{ن}{٢}$ ، $\frac{ن}{٢}$ + ١ ثم نحسب متوسط قيمتهما فتكون هـــي وســـيط البيان.

مثال (٢) :

نفترض أنه لدينا شركة عاشرة رأس مالها ١٢ مليون جنيه أضـــيفت إلـــى الشركات بالمثال (١) فإن ترتيب المفردات يصبح كما هو سابقاً مـــع إضـــافة القيمة ١٢ في نهاية الترتيب.

٣,٥ ، ٤ ، ٥ ، ٦,٥ ، [٧ ، ٨] ، ٨,٧ ، ٩,٥ ، ١٠ ، ١٢

موقع المفردتين التي تجعل عدد المفردات قبلهما مساوياً لعـــدد المفـــردات بعدهما هما :

$$\frac{ن}{٢} = \frac{١٠}{٢} = ٥ \text{ ، } \frac{ن}{٢} + ١ = \frac{١٠}{٢} + ١ = ٦$$

∴ موقع المفردتين هما رقم ٥، رقم ٦ في الترتيب أعلاه وقيمتهما ٧ ، ٨ وبالتالي فإن وسيط رأس المال = $\frac{٧ + ٨}{٢}$ = ٧,٥ مليون جنيه.

لكن ما يهمنا أكثر هو كيفية حساب الوسيط من جداول التوزيع التكراريــة. وما سنلاحظه أن حساب الوسيط لا يتأثر بهل الفئات منتظمة أم غير منتظمــة، مفتوحة أم غير مفتوحة، لذلك يعتبر المقاييس المميزة بأشياء غير متوافرة فــي مقاييس أخرى للمتوسطات.

كيفية إيجاد الوسيط من جداول تكرارية :

مثال (٣) :

الجدول التالي يبين توزيع ١٢٠ طالباً حسب فئات العمر المختلفة.

فئات العمر	أقل من ١٠	١٠ –	١٢–	١٥–	١٨–	٢٠–٢٤	المجموع
عدد الطلاب	٤	٢٠	٢٨	٤٠	٢٣	٥	١٢٠

يعتبر الوسيط من المقاييس الإحصائية للمتوسطات التي يمكن إيجادها بالحساب والرسم.

إيجاد الوسيط حسابياً :

لحساب الوسيط للعمر نبدأ بالخطوات التالية :

١) عمل جدول متجمع صاعداً أو هابط. وهذه الخطوة هي التي تقوم مقام الترتيب التصاعدي أو التنازلي للبيانات.

نتائج تكوين الجدولين المتجمع الصاعد والهابط كما يلي :

جدول (٢)

ف	ك	حدود عليا للفئات	ك متجمع صاعد
أقل من ١٠	٤	أقل من ١٠	٤
١٠ – ١٢	٢٠	أقل من ١٢	٢٤
١٢ – ١٥	٢٨	أقل من ١٥	٥٢
١٥ – ١٨	٤٠	أقل من ١٨	٩٣
١٨ – ٢٠	٢٣	أقل من ٢٠	١١٥
٢٠– ٢٤	٥	أقل من ٢٤	١٢٠
المجموع	١٢٠ مجــ ك		

جدول (٣)

ف	ك	حدود دنيا للفئات	ك متجمعة هابطة
أقل من ١٠	٤	الحد الأدنى فأكثر	١٢٠
١٠ – ١٢	٢٠	١٠ فأكثر	١١٦
١٢ – ١٥	٢٨	١٢ فأكثر	٩٦
١٥ – ١٨	٤٠	١٥ فأكثر	٦٨
١٨ – ٢٠	٢٣	١٨ فأكثر	٢٨
٢٠– ٢٤	٥	٢٠ فأكثر	٥
المجموع	١٢٠ مجــ ك		

٢) ترتيب الوسيط = $\frac{\text{مجـ ك}}{٢} = \frac{١٢٠}{٢} = ٦٠$

٣) بالنظر إلى العمود ك متجمع صاعد نجد أن ترتيب الوسيط ٦٠ يقع ما بين ٥٢، ٩٢ مما يجعل قيمة الوسيط تقع ما بين الحدين المقابلين وهما ١٥ ، ١٨. هذه الخطوة يعاد ترتيبها كما هي في الجدول المتجمع الصاعد مع كتابة دلالات للنواتج الأربعة التي أخذت منه :

١٥	٥٢	تكرار سابق
	٦٠	ترتيب الوسيط
١٨	٩٢	تكرار لاحق

الفئة الوسيطة التي سيقع فيها الوسيط حيث بدايتها ١٥ ونهايتها ١٨ وطولها ٣.

بتطبيق القانون التالي :

$$\text{قيمة الوسيط} = \text{بداية فئة الوسيط} + \frac{\text{ترتيب الوسيط} - \text{ك سابق}}{\text{ك لاحق} - \text{ك سابق}} \times \text{طول فئة الوسيط}$$

$$= ١٥ + \frac{٦٠ - ٥٢}{٩٢ - ٥٢} \times ٣$$

$$\text{قيمة الوسيط} = ١٥ + \frac{٨}{٤٠} \times ٣ = ١٥ + ٠,٦ = ١٥,٦$$

يلاحظ أننا بدأنا مثالنا السابق بفئات غير منتظمة وكذلك مفتوحة من أعلــى ونفس الشيء إذا كانت مفتوحة من أسفل أو من الطرفين ولم تمثل هذه الحالات أي قيد على خطوات الحل فمن السهل إيجاد الوسيط بــنفس الطريقــة للفئــات المنتظمة والمغلقة.

فإذا نظرنا إلى عمود ك متجمع هابط فإننا نجد أن ترتيب الوسيط ٦٠ يقــع ما بين ٦٨ ، ٢٨ مما يعل قيمة الوسيط ما بين حدين مقــابلين همــا ١٥ ، ١٨ بالجدول ويكون تنظيم بيانات هذه الخطوة كما هو مبين.

فئة الوسيط				
	١٥		٦٨	ك سابق
فئة الوسيط		قيمة الوسيط		٦٠ ترتيب الوسيط
	١٨		٢٨	ك لاحق

بتطبيق نفس القانون :

$$\therefore \text{قيمة الوسيط} = ١٥ + \frac{٦٠ - ٦٨}{٢٨ - ٦٨} \times ٣ = ١٥,٦$$

ويعتبر هذا عرضاً لإيجاد الوسيط حسابياً بطريقتي الجدول المتجمع الصاعد والجدول المتجمع الهابط. لكن إذا طلب منا هذا المقياس فيكفي أســلوب واحــد لتحديد المقياس.

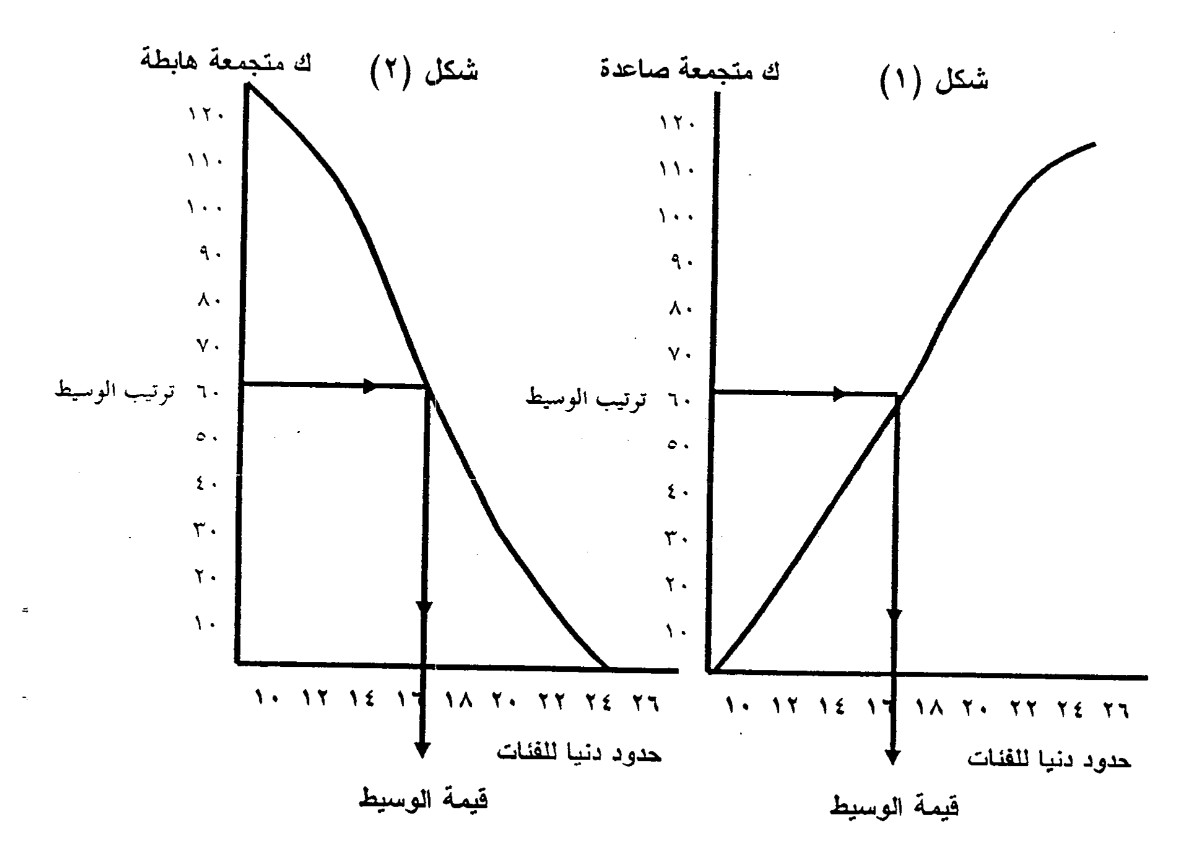
إيجاد الوسيط بالرسم :

تعتبر الخطوتان الأولى والثانية لإيجاد الوسيط حســابياً أيضــاً لاســتنباط الوسيط بالرسم فتصبح الخطوات هي :

١) عمل الجدول المتجمع الصاعد أو الهابط ويمثلهما جـدولي (٢) ، (٣) السابقين.

٢) إيجاد ترتيب الوسيط = مجـ ك ÷ ٢ = ١٢٠ ÷ ٢ = ٦٠

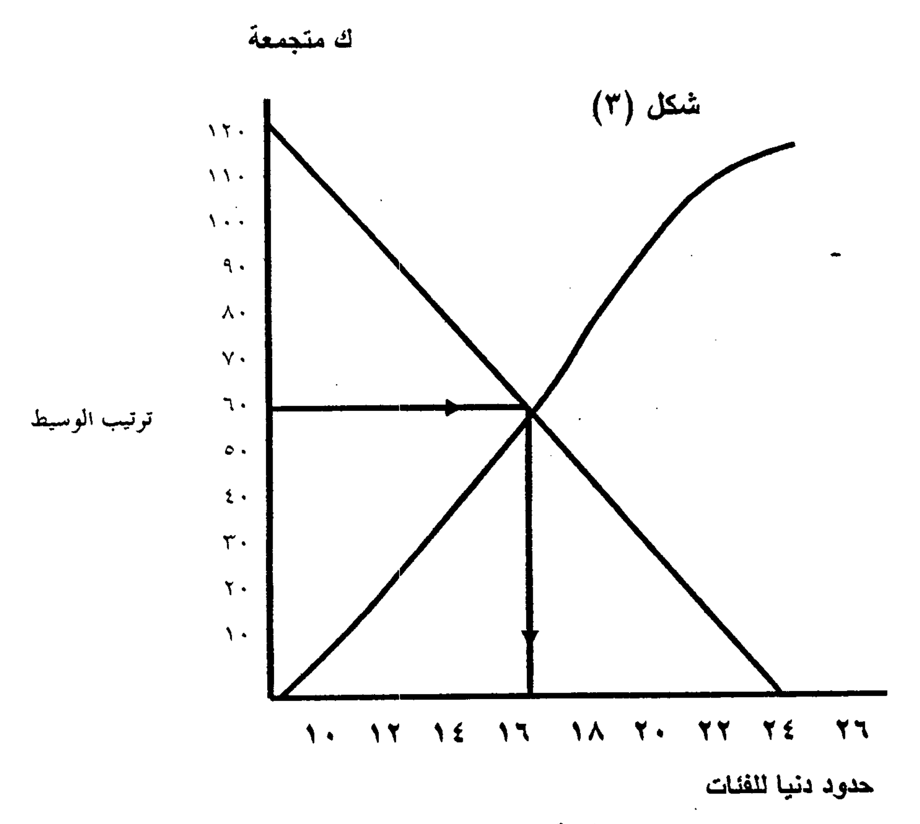
٣) نرسم الجدول المتجمع الصاعد أو الهابط بالمنحنى المتجمع الصاعد أو الهابط. وعلى هذا الأساس يظهر المنحنيين المطلوبين في مثالنا هذا في الشكلين الآتيين :



٤) عن طريق الخطوتين الأخيرتين تحدد ترتيب الوسيط ٦٠ على المحور الرأسي (ك متجمعة الصاعدة أو الهابطة) ثم نقيم عمود عند الترتيب ليقابل المنحنى في نقطة ثم نسقط من هذه النقطة عموداً على المحور الأفقي ليقابل المنحنى في نقطة هي في الواقع قيمة الوسيط المطلوب.

والجدير بالذكر أن المنحنيين الصاعد والهابط يمكن رسمهما معاً لنفس مقياس الرسم كما في شكل (٣). فإذا أسقطنا عمود نقطة تلاقي المنحنيين على المحور الأفقي يتحدد عندها قيمة الوسيط أما العمود الساقط على المحور الرأسي فيتحدد عنده ترتيب الوسيط.

وتجدر الإشارة أيضاً إلى أن القيمة المستخرجة من الرسم غالباً ما تكون قيمة تقريبية للقيمة التي نحصل عليها بطريق الحساب حيث تمهيد المنحنى وطرق التدريج للمحورين الأفقي والرأسي ورصد النقاط يتم بالطرق البدوية.



٣) المنوال Mode :

يعتبر المنوال أحد مقاييس النزعة المركزية للبيانات الإحصائية. ويمكن تعريفه على أنه تلك القيمة الأكثر شيوعاً أو التي تتكرر أكثر من غيرها من قيم الظاهرة. ويمكن حساب قيمة المنوال سواء من البيانات غير المبوبة أو تلك المبوبة في جداول توزيع تكراري. وسنتناول فيما يلي كيفية حسابه.

١. حساب المنوال من البيانات غير المبوبة :

إذا كانت درجات امتحان مادة الإحصاء لستة طلاب هي :

65, 70, 80, 70, 85, 70

فإن المنوال لتلك المجموعة هو القيمة الأكبر تكراراً :

Mo = 70

أما إذا كانت درجات هؤلاء الطلاب هي :

65, 70, 80, 83, 85, 81

فإنه لا يوجد منوال لتلك المجموعة.

٢. حساب البيانات من البيانات المبوبة :

يمكن إيجاد قيمة المنوال إما بالطريقة الحسابية أو بالطريقة البيانية، وتجدر الإشارة إلى أن إتباع أي من الطريقتين يتطلب الخطوات المبدئية التالية :

١. تعديل تكرارات الفئات إذا كانت الفئات غير متساوية.
٢. تحديد الفئة المنوالية وهي الفئة الأكثر تكراراً، على أن يتم ذلك بعد إجراء تعديل الفئات غير المتساوية إن وجدت لأنه ليس بالضرورة أن تكون الفئة المنوالية قبل التعديل هي نفسها بعد التعديل.

أ) الطريقة الحسابية :

هناك طرق عديدة لحساب قيمة المنوال، غير أننا سنقتصر على دراسة طريقة بيرسون فقط باعتبارها من أهم الطرق المتبعة ويمكن إيجاد قيمة المنوال وفقاً للمعادلة التالية :

$$\mathbf{Mo = L_{Mo} + \frac{d_1}{d_1 + d_2} i_{Mo}}$$

حيث :

الحد الأدنى للفئة المنوالية $= L_{Mo}$

الفرق المطلق بين تكرار الفئة المنوالية والفئة السابقة لها $= d_1$

الفرق المطلق بين تكرار الفئة المنوالية والفئة التالية لها $= d_2$

طول الفئة المنوالية $= i_{Mo}$

مثال : الجدول التالي يبين توزيع ٤٥ طالباً من طلاب المعهد وفقاً لـــدرجاتـهم في امتحان مادة الإحصاء

Glass	: 30-	40-	50-	60-	70-	80-	90-100
Frequency	: 6	7	4	[16]	4	5	3

والمطلوب حساب قيمة المنوال لهذا التوزيع التكراري.

الحل :

من الواضح أن التوزيع التكراري منتظم وأن أكبر تكرار هو (16) وبذلك تكون الفئة المنوالية هي : (60-70).

$L_{Mo} = 60 \qquad d_1 = 12 \qquad d_2 = 12 \qquad i_{Mo} = 10$

$$Mo = 60 + \frac{12}{12 + 12} \times 10$$

$$= 60 + 5 = 65$$

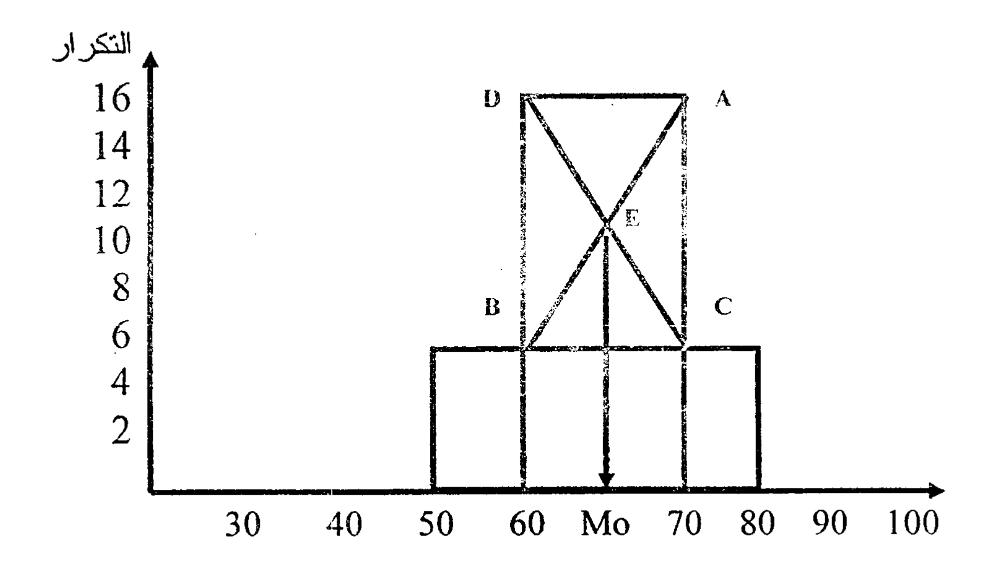
$$Mo = 65$$

ب) **الطريقة البيانية :**

يمكن إيجاد المنوال بالاستعانة بالمدرج التكراري، حيث يكتفي بالتوقيع البياني لكل من الفئة المنوالية والفئتين السابقة عليها واللاحقة لها. ويجب أن يؤخذ في الاعتبار تعديل التكرارات أولاً في حالات التوزيعات التكرارية غير المنتظمة.

مثال : مستعيناً بالتوزيع التكراري السابق اشتق قيمة المنوال بيانياً.

الحل : نقوم برسم المدرج التكراري للفئات الثلاث المذكورة ومنه نشتق المنوال على النحو التالي :



ومن الشكل يتبين كيفية الاشتقاق، حيث يتم توصيل المستقيمين AB ، CA وعند نقطة تقاطعها (E) نسقط عموداً على المحور الأفقي فيحدد قيمة المنوال، وكما هو واضح من الشكل فإن :

$Mo = 65$

وهي نفس القيمة التي سبق التوصل إليها بالطريقة الحسابية.

خصائص المنوال :

١) لا يأخذ في حسابه جميع مفردات الظاهرة، ومن ثم لا يتأثر القيم المتطرفة.

٢) يمكن حسابه من الجداول التكرارية المفتوحة من أحد طرفيها أو من كليهما.

٣) سهولة حسابه، إذ يحتاج فقط إلى معرفة القيمة الأكثر تكراراً من غيرها في التوزيع.

٤) يتطلب حسابه في حالة التوزيعات التكرارية غير المنتظمة ضرورة تعديل التكرارات.

٥) تتأثر قيمة المنوال بطريقة ترتيب مفردات التوزيع التكـراري، حيـث أن تعديل حدود الفئات يترتب عليه تغيير توزيع التكرارات على هذه الفئـات، وبالتالي تتغير قيمة المنوال.

٦) قد لا يوجد منوال لبعض الظواهر، كما أنه لا يخضع للعمليات الجبرية كما هو الحال بالنسبة للوسط الحسابي.

٤) الوسط الهندسي Geometric Mean :

الوسط الهندسي G لمجموعة من N رقم $X_1, X_2, X_3, ..., X_N$ هـو الجـذر النوني لحاصل ضرب هذه الأرقام.

ويعتبر الوسط الهندسي متوسطاً رياضياً كالوسط الحسابي وليس متوسـطاً مكانياً كالوسيط. ويفضل استخدامه عند تقدير عدد السكان بين فتـرة التعـداد، وعند حساب متوسطات النسب أو المعدلات.

ويمكن حساب الوسط الهندسي من البيانات الموزعة تكرارياً (المبوبـــة) أو الغير موزعة تكرارياً (غير مبوبة).

حساب الوسط الهندسي من البيانات الغير مبوبة :

إذا كان لدينا عدد من القيم $X_1, X_2, X_3, \ldots, X_N$ ورمزنا للوســط الهندســي بالرمز G فإن :

$$G = \sqrt[N]{X_1, X_2 X_3 \ldots\ldots X_N}$$

ويأخذ لوغريتمات الطرفين فإن :

$$\text{LogG} = \frac{1}{N}\left(\text{LogX}_1 + \text{LogX}_2 + \text{LogX}_3 + \ldots + \text{LogX}_N\right)$$

$$= \frac{1}{N}\left(\sum \text{LogX}\right) = \frac{\sum \text{LogX}}{N}$$

مثال : احسب الوسط الهندسي للقيم الآتية :

$$14, 11, 22, 23, 5, 6$$

$$\text{LogG} = \frac{1}{6}\left(\text{Log14} + \text{Log11} + \text{Log22} + \text{Log23} + \text{Log5} + \text{Log6}\right)$$

$$= \frac{1}{6}\left(1.146 + 1.0414 + 1.3424 + 1.3617 + 0.6990 + 0.7781\right)$$

$$= \frac{1}{6}\left(6.3687\right) = 1.06145$$

$$\therefore G = 11.5201$$

حساب الوسط الهندسي من البيانات الموزعة تكرارياً (المبوبة) :

$$G = N\sqrt{X_1^{F_1} X_2^{F_2} \ldots\ldots X_K^{F_K}}$$

وبأخذ لوغريتمات الطرفين فإن :

$$\mathbf{LogG = \frac{1}{N}(F_1LogX_1 + F_2LogX + ... + F_K LogX_K)}$$

$$\mathbf{= \frac{1}{N}\sum_{J=1}^{K} F_j LogX_j = \frac{\sum FLogX}{N}}$$

حيث : $\mathbf{N = \sum F}$

مثال : أحسب الوسيط الهندسي للبيانات في الجدول التكراري الآتي:

الفئات	التكرارات F	مراكز الفئات X	لوغـــــاريتم مراكز الفئات LogX	FLogX
8 –	5	9	0.9542	7.6336
10 –	12	11	1.0414	12.4968
12 –	14	13	1.1139	15.5946
14 –	8	15	1.1761	9.4088
16 - 18	6	17	1.2304	7.3824
∑	48			

$$\mathbf{LogG = \frac{\sum FLogX}{N}}$$

$$\mathbf{= \frac{52.5162}{48} = 1.09409}$$

$$\mathbf{\therefore G = 12.419}$$

والوسط الهندسي لا يمكن استعماله إذا كانت قيمة أو أكثر من القيم تســاوي الصفر أو سالبة، وذلك لأن ضرب القيم في بعضها يجعل الناتج صفراً، وكذلك

لأن جذر الأرقام السالبة عدد غير حقيقي، ولذلك فهو غير شـــائع الاســـتعمال لصعوبة حسابه وضرورة استخدام جداول اللوغاريتمات.

أهم خواص الوسط الهندسي :

١. تعريفه محدد وواضح وقيمته وحيدة.

٢. يحتاج حسابه لعمليات حسابية معقدة.

٣. لا يمكن حسابه إذا كانت إحدى القيم صفراً أو سالبة.

٤. لا يمكن حسابه من جداول مفتوحة.

٥) الوسط التوافقي Harmonic Mean

الوسط التوافقي لمجموعة من القيم هو مقلوب الوسط الحسابي لمقلوب هـــذه القيم.

حساب الوسط التوافقي من بيانات غير مبوبة :

$$H = \frac{N}{\sum\left(\frac{1}{X}\right)}$$

مثال : احسب الوسط التوافقي للبيانات التالية :

6, 15, 11, 14, 22, 13

$$H = \frac{6}{\left(\frac{1}{6}+\frac{1}{15}+\frac{1}{11}+\frac{1}{14}+\frac{1}{22}+\frac{1}{13}\right)}$$

$$= \frac{6}{(0.1667+0.0667+0.0909+0.0714+0.0455+0.0769)}$$

$$= \frac{6}{0.5181} = 11.5808$$

مثال : ثلاث سيارات سرعة الأولى 80km/h وسرعة الثانيـــة 110km/h وسرعة الثالثة 40km/h احسب متوسط سرعة السيارات الثلاث ؟

يلاحظ أن حساب الوسط الحسابي في هذه الحالـــة غيـــر دقيـــق لأن هـــذه معدلات، ومن ثم فإن أنسب معيار للحساب هو الوسط التوافقي.

$$H = \frac{3}{\left(\frac{1}{80} + \frac{1}{110} + \frac{1}{40}\right)}$$

$$= \frac{3}{(0.0125 + 0.0091 + 0.025)}$$

$$H = \frac{3}{0.0466} = 64.3777 km/h$$

وإذا قارنا هذه النتيجة بالوسط الحسابي الذي قيمته

$$\overline{X} = \frac{80 + 110 + 40}{3} = \frac{230}{3} = 67.6667 km/h$$

فإن هذا يوضح مدى خطأ استخدام الوسط الحسابي في تقـــدير متوســـطات السرعة.

حساب الوسط التوافقي من البيانات المبوبة :

$$H = \frac{\sum F}{\sum\left(\frac{F}{X}\right)}$$

مثال : احسب الوسط التوافقي من البيانات في الجدول التكراري التالي :

الفئات	التكرارات F	X	$\frac{1}{X}$	$F\left(\frac{1}{X}\right)$
8 –	8	9	0.11111	0.88888
10 –	12	11	0.09090	1.09090
12 –	14	13	0.07692	1.07688
14 –	8	15	0.06666	0.53328
16-18	6	17	0.05882	0.35292
$\sum$	48			

$$H = \frac{48}{3.94286} = 12.173$$

أهم خواص الوسط التوافقي :

١. يحتاج حسابه لعمليات حسابية معقدة.

٢. لا يمكن حسابه من جداول مفتوحة.

٣. أفضل المتوسطات في حالة المعدلات (معدل التغير – معدلات السرعات)

٤. لا يمكن حسابه إذا كانت إحدى القيم صفراً (لأن مقلوب الصفر كمية غير معينة) أو إذا كان مجموع مقلوبات القيم يساوي صفراً.

ملاحظات :

١. الوسط الحسابي لأي مجموعة من البيانات أكبر من الوسط الهندسي، والوسط الهندسي أكبر من الوسط التوافقي أي أن :

الوسط الحسابي > الوسط الهندسي > الوسط التوافقي

مثال : احسب الوسط الحسابي والهندسي والتوافقي لمجموعة البيانات التالية :

11, 22, 14, 15, 10

$$\overline{X} = \frac{72}{5} = 14.4$$

$$\text{LogG} = \frac{\sum \text{LogX}}{N}$$

$$= \frac{5.706034}{5} = 1.14120694$$

$$\therefore G = 13.842258$$

$$H = \frac{N}{\sum(\frac{1}{X})}$$

$$H = \frac{5}{\left(\frac{1}{11} + \frac{1}{22} + \frac{1}{14} + \frac{1}{15} + \frac{1}{10}\right)}$$

$$= \frac{5}{0.3744587} = 13.3526$$

يلاحظ من هذا المثال أن :

$$\overline{X} > G > H$$

٢. الوسط الهندسي والتوافقي يمتازان عن الوسط الحسابي بأنهما أقل تأثراً بالقيم الشاذة، والوسط الهندسي يفضل في حالة حساب متوسطات النسب.

٣. يعاب على الوسط الحسابي والهندسي والتوافقي بأنه من الصعب حسابهم من جداول التوزيع التكراري المفتوحة.

٦) العلاقة بين الوسط الحسابي والوسيط والمنوال :

توجد علاقة بين الوسط والوسيط والمنوال تتضح في :

الوسط الحسابي = الوسيط = المنوال

وذلك عندما يكون التوزيع متماثل، أي أن :

$\overline{X} = Md = Mo$

وفي حالة التوزيع الغير متماثل :

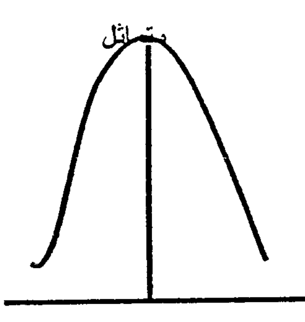
الوسط الحسابي ـــ المنوال = ٣ (الوسط الحسابي ـــ الوسيط)

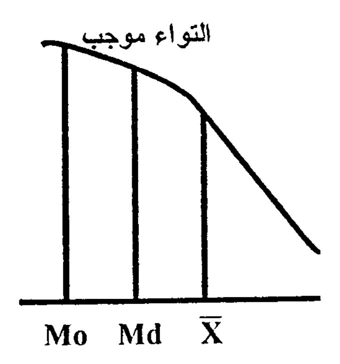
في حالة الالتواء الموجب يكون :

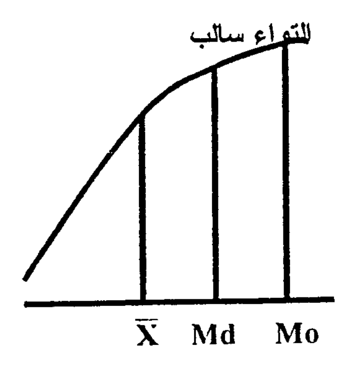
المنوال < الوسيط < الوسط الحسابي

وفي حالة الالتواء السالب يكون :

النوال > الوسيط > الوسط الحسابي







$\overline{X}$

$= Md$

$= Mo$

الفصل السادس

مقاييس التشتت

Measures of Dispprsion

لا تكفي المتوسطات وحدها لإعطاء فكرة دقيقة عن المجموعـة العدديـة، فالمجموعتان الآتيتان :

24, 16, 13, 7, 5
15, 14, 13, 12, 11

مشتركتان في الوسط الحسابي (13) والوسيط (13)، رغم تفاوتهما الشاسع في توزيع مفرداتهما، فمفردات المجموعة الأولى منتشرة في مدى قدرة (19)، بينهما مفردات المجموعة الثانية منتشرة في مدى قدرة (4).

ومقاييس التشتت هي المقاييس التي تقيس مدى القرب أو البعد عن القيمـة المتوسطة،ويمكن تقسيم مقاييس التشتت إلى نوعين : النوع الأول هو (مقاييس التشتت المطلقة)، والنوع الثاني هو (مقاييس التشتت النسبية)، والفـرق بـين النوعين هو أن النوع الأول يعطي مقياس لتشتت البيانات ويأخذ المقياس وحدة قياس هي نفس وحدة قياس القيم الأصلية، فإذا كانت بالجنيه سـتكون النتيجـة بالجنيه وإذا كانت بالمتر ستكون النتيجة بالمتر ... الخ، أما مقـاييس التشـتت النسبية فليس لها تمييز، وتمتاز عن المطلقة في أنهـا تصـلح للمقارنـة بـين مجموعتين إذا كانت وحدات القياس في المجموعتين مختلفتين.

وفي هذا الفصل سنتناول بالمناقشة بعض مقاييس التشتت المطلقة وهي :

المدى ونصف المدى الربيعي
RANGE & SEMI INTERQUARTILE RANGE

يعرف المدى بأنه الفرق بين أكبر قيمة في التوزيع وأصغر قيمة . ويمكن حساب المدى من البيانات الغير مبوبة والمبوبة :

حساب المدى من البيانات غير مبوبة :

مدى مجموعة من الأرقام هو الفرق بين أكبر وأقل رقم في المجموعة.

مثال : مدى المجموعة 12 , 10 ,8 ,5 ,5 ,5 ,3 ,3 ,2 هو (10) .

حساب المدى من البيانات المبوبة :

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى .

أو المدى = مركز الفئة الأخيرة – مركز الفئة الأولى .

مثال : أحسب المدى لجدول التوزيع التكراري الآتي :

الفئات	التكرارات F	مراكز الفئات X
8-	8	9
10-	12	11
12-	14	13
14-	8	15
16-18	6	17

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

$$\text{Range} = 18 - 8 = 10$$

أو المدى = مركز الفئة الأخيرة – مركز الفئة الأولى

$$\text{Range} = 17 - 9 = 8$$

ويكون المدى في الحالة الأخيرة مساوياً للمدى في الحالة الأولى وذلك عندما يُضاف طول الفئة $8 + 2 = 10$ وذلك عندما تكون الفئات متساوية الأطوال .

خواص ومميزات المدى :

١- يمتاز المدى ببساطته وسهولة تفهم معناه .

٢- سهل في حسابه لأن قيمته تعتمد على الفرق بين قيمتين فقط .

عيوبـــــه :

١- أقل مقاييس التشتت كفاءة لأن قيمته تعتمد فقط على الفرق بين أعلـــى وأقل قيمة في التوزيع ، لذلك فقد تصبح قيمته شـــاذة إذا كانـــت تلـــك القيمتين من القيم المتطرفة أو الشاذة .

٢- يُحسب المدى في الجداول التكرارية عن طريق الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى ، وبالتالي لا يمكن حسابه مـــن جداول مفتوحة .

وللتغلب على مشكلة تأثر المدى بالقيم الشاذة (المتطرفة) يمكننا تعريف مقياس آخر يعتمد على إهمال جزء من البيانات عند طرفـــي التوزيـــع حتى نتخلص من القيم الشاذة ، هذا المقياس يُطلق عليه اســـم نصـــف المدى الربيعي ويُحسب كالآتـــي :

نصف المدى الربيعي = (الربيع الأعلى – الربيع الأدنـــى) / ٢

مثال : إحسب نصف المدى الربيعي للأطوال الآتيـــة

162, 154, 174, 158, 180, 182, 186

١ – نرتب القيم تصاعدياً أو تنازلياً :

154, 158, 162, 174, 179, 180, 182, 186

٢ – نحسب ترتيب الربيع الأول :

$$\frac{N+1}{4} = \frac{9+1}{4} = 2.5$$

٣ – نحسب ترتيب الربيع الثالث :

$$\frac{3(N+1)}{4} = \frac{3(9+1)}{4} = 7.5$$

٤ – نوجد قيمة الربيع الأول $158 + \frac{1}{2}(162-158) = 158 + 2 = 160$

٥ – نوجد قيمة الربيع الثالث $180 + \frac{1}{2}(182-180) = 180 + 1 = 181$

$$Q = \frac{181 - 160}{2} = \frac{21}{2} = 15.5$$

مثال : إحسب نصف المدى الربيعي من الجدول التالي :

الفئات	التكرارات	الحدود العليا للفئات	التكرار المتجمع الصاعد
0 -	15	أقل من 4	15
4 -	35	أقل من 8	50
8 -	60	أقل من 12	110
12 -	60	أقل من 16	170
16 -	30	أقل من 20	200

ترتيب الربيع الأول = $\frac{200}{4} = 50$

ترتيب الربيع الثالث = $\frac{200 \times 3}{4} = 150$

قيمة الربيع الأول = $4 + \frac{50 - 15}{50 - x15} \times 4 = 4 + 4 = 8$

قيمة الربيع الثالث = $12 + \frac{150 - 110}{170 - 110} \times 4 = 12 + 2.67 = 14.7$

نصف المدى الربيعي = $\frac{14.7 - 8}{2} = 6.7 = 3.3$

الإنحراف المتوســــــــط

AVERAGE (MEAN) DEVIATION

إنه من المهم أن نقيس تشتت المجموعــة حــول أحــد متوســطاتها ، وخصوصاً وسطها الحسابي ، وفي هذه الحالة نعتبر قيمــة الوســط الحســابي كمركز ، ونوجد انحرافات قيم مفردات المجموعة عن هذا الوســط الحســابي باستخدام العلاقة التالية :

$$M.D. = \frac{\Sigma (X - \bar{X})}{N}$$

حيث $| x - \bar{x} |$ هو الانحراف العددي لقيم مفردات المجموعة عن الوســط الحسابي $\bar{x}$ ، أي بدون الإشارات + ، لأن التشتت معنــاه البعــد عــن x دون التمييز بين النقص والزيادة .

مثال : إحسب الانحراف المتوسط لمجموعة البيانات التالية :

دون التمييز بين النقص والزيادة .

مثال : إحسب الانحراف المتوسط لمجموعة البيانات التالية :

37, 45, 42 , 41, 51

x	$\lvert X - \bar{x} \rvert$
37	6.2
45	1.8
42	1.2
41	2.2
51	7.8
216	19.2

$$M.D. = \frac{19\text{-}2}{5} = 3.84$$

أما في حالة البيانات التكرارية فإن $M.D. = \frac{\Sigma f|x-\bar{x}|}{N}$

حيث $N = \Sigma f$

مثال : إحسب الانحراف المتوسط من الجدول الآتـــي :

فئات	F	x	$X - \bar{x}$	$F(x - \bar{x})$	x-md	F(x – md)
0 –	5	5	13	65	11.25	56.25
10-	8	15	3	24	1.25	10.00
20-	3	25	7	21	8.75	26.25
30-40	4	35	17	68	18.75	75.00
Σ	20			178		167.50

وبذلك يكون :

الانحراف المتوسط عن المتوسط

$$M.D(\bar{x}) = \frac{178}{20} = 8.9$$

الانحراف المتوسط عن الوسيط

$$M.D(md) = \frac{167.5}{20} = 8.375$$

يعاب على الانحراف المتوسط أنه لا يمكن حسابه من جداول تكرارية مفتوحة .

الانحراف المعياري

STANDERD DEVIATION

الانحراف المعياري (S) هو أهم وأدق مقاييس التشتت المعروفة حول الوسط الحسابي ($\overline{X}$) وأكثرها استخداماً في علم الإحصاء .

تعريفه : هو الجذر التربيعي لمتوسط مربعات انحرافات قيم مفردات المجموعة عن الوسط الحسابي ($\overline{X}$) .

حساب الانحراف المعياري من البيانات الغير (مبوبة) :

هناك ثلاث طرق لحساب الانحراف المعياري ....

الطريقة الأولى : وهي الطريقة المطولة :

$$S = \sqrt{\frac{\Sigma x^2}{n} - \left\{\frac{\Sigma x}{n}\right\}^2}$$

الطريقة الثانية – وهي الطريقة المختصرة :

$$S = \sqrt{\frac{\Sigma d^2}{n} - \left\{\frac{\Sigma d}{n}\right\}^2}$$

الطريقة الثالثة – وهي الطريقة الأكثر اختصاراً :

$$S = c\sqrt{\frac{\Sigma u^2}{n} - \left\{\frac{\Sigma u}{n}\right\}^2}$$

مثال : إحسب الانحراف المعياري من البيانات الآتية :

37, 45, 42, 41, 51

X	x^2	d x-A	d^2	U d/c	U^2
37	1369	-5	25	-1.67	2.79
45	2025	3	99	1.00	1.00
42	1764	0	0	0	0
41	1681	-1	1	-0.33	0.11
51	2601	9	81	3.00	9.00
216	9440	6	116	2	12.90

الطريقة المطولـــة

$$\sqrt{\frac{9440}{5} - \left\{\frac{216}{5}\right\}^2}$$

$$\sqrt{1888 - 1866.24}$$

$$\sqrt{21.76} = 4.66$$

الطريقة المختصرة

$$\sqrt{\frac{116}{5} - \left\{\frac{6}{5}\right\}^2}$$

$$\sqrt{23.2 - 1.44}$$

$$\sqrt{21.76} = 4.66$$

الطريقة الأكثر اختصارا

$$3\sqrt{12.90 - \left\{\frac{2}{5}\right\}^2}$$

$$3\sqrt{2.58 - 0.16}$$

$$3\sqrt{2.42}$$

$$3 \times 1.555 = 4.66$$

حساب الانحراف المعياري من البيانات المبوبة :

مثال : احسب الانحراف المعياري من الجدول التكراري التالي :

فئات	f	A			B			c		
		X	fx	Fx²	d	fd	Fd²	u	fu	Fu²
0-	5	5	25	125	-10	-50	500	-1	-5	5
10-	8	15	120	1800	0	0	0	0	0	0
20-	3	25	75	1875	10	300	300	1	3	3
30-40	4	35	140	4900	20	80	1600	2	8	16
	20		360	8700		60	2400		6	24

يُحسب الانحراف المعياري بالطريقة المختصرة من الجزء (B) من الجـدول ، ولقد تم الحصول على العمود الأول في الجزء (B) من مراكز الفئات بعد طرح مقدار ثابت وليكن 10 .

ويمكن حساب الانحراف المعياري من المعادلة الآتية :

$$S = \sqrt{\frac{\Sigma\ f\ d\ 2}{\Sigma f} - \left\{\frac{\Sigma\ f\ d}{\Sigma f}\right\}^2}$$

$$= \sqrt{\frac{2400}{20} - \left\{\frac{60}{20}\right\}^2}$$

$$= \sqrt{120 - 9} = \sqrt{111} = 10.536$$

معامل الاختــــلاف
COEFFICIENT OF VARIATION

معامل الاختلاف (C.V) يعبر عن الانحراف المعياري فـــي صـــورة نسبة مئوية من الوسط الحسابي . وعلى ذلك يمكن استخدامه في مقارنة التشتت للمجموعات المختلفة حتى ولو لم تتساوى أوساطها الحسابية ، أو اختلاف نوع الوحدات المستعملة في قياس مفرداتها .

ويمكن حساب معامل الاختلاف بعدة طرق :

الطريقة الأولى – باستخدام الوسط الحسابي والانحراف المعياري :

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times ١٠٠$$

الطريقة الثانية – باستخدام الربيعين الأعلى والأدنى :

$$\text{معامل الاختلاف} = \frac{\text{الربيع الأعلى – الربيع الأدنى}}{\text{الربيع الأعلى + الربيع الأدنى}} \times ١٠٠$$

الطريقة الثالثة – باستخدام الربيعين والوسيط :

$$\text{معامل الاختلاف} = \frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{\text{٢ الوسيط}} \times ١٠٠$$

ويفضل استخدام أي من الطريقتين الثانية والثالثة عندما تكون الجداول التكرارية مفتوحة .

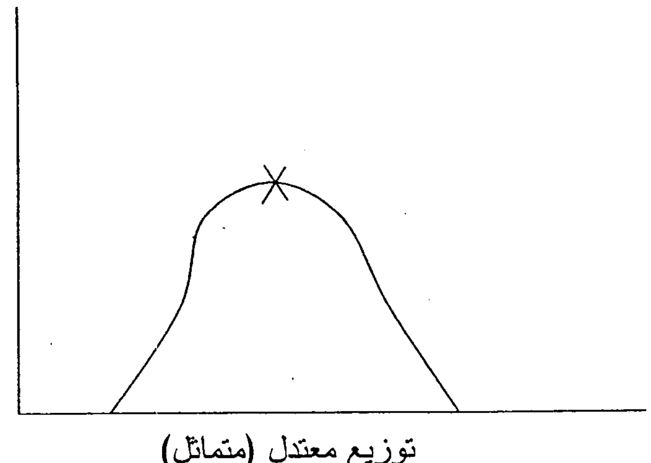
الفصل السابع

معامل الالتواء

Sekw Ness

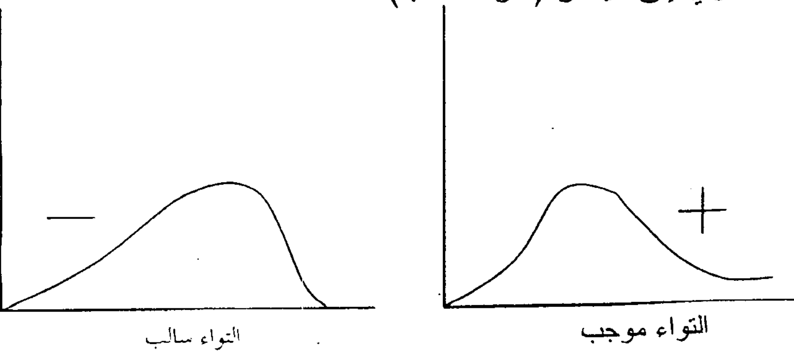
ذكرنا في المتوسطات أنه في حالة التوزيعات التكرارية المتماثلة فإننا نجد أن :

الوسيط الحسابي = الوسيط = المنوال



توزيع معتدل (متماثل)

ولكن قد تكون التوزيعات التكرارية غير متماثلة أي نجد أن المنحنى يتجه إلى اليمين أو يتجه جهة اليسار . فنقول أن التوزيع ملتوي إلى اليمين (التواء موجب) أو ملتوي إلى اليسار (التواء سالب) .



وترجع أهمية الالتواء إلى أنه في بعض الحالات قد نجد بالنسبة لتوزيعين تكرارين مختلفين أن القيمة الوسيطية لهما واحدة وكذلك قيمة التشتت واحدة في حين أن التواء كل منهما يكون مختلف عن الآخر .

وهناك مقاييس للالتواء منها :

$$\text{معامل الالتواء} = \frac{\text{الوسط الحسابي} - \text{المنوال}}{\text{الانحراف المعياري}} \qquad (١)$$

$$\text{معامل الالتواء} = \frac{٣\ (\text{الوسط الحسابي} - \text{الوسيط})}{\text{الانحراف المعياري}} \qquad (٢)$$

معامل الالتواء =

$$\frac{(\text{قيمة الربيع الأعلي} - \text{الوسيط}) - (\text{الوسيط} - \text{قيمة الربيع الأدنى})}{(\text{قيمة الربيع الأعلى} - \text{الوسيط}) + (\text{الوسيط} - \text{قيمة الربيع الأدنى})}$$

والقانون رقم ١ ، ٢ لا يمكن إيجاده بالرسم حيث لا يمكن إيجاد قيمة الوسط الحسابي أو الانحراف المعياري بالرسم كذلك لا يمكن استخدام هذين القانونين في حالة التوزيعات التكرارية المفتوحة بل يشترط لاستخدامها أن تكون التوزيعات التكرارية مقفولة .

أما القانون رقم ٣ فيمكن إيجاده بالرسم البياني حيث يمكن إيجاد قيمة كل من الربيع الأعلى والأدنى وقيمة الوسيط برسم المنحنى المتجمع الصاعد أو المنحنى المتجمع الهابط واستنتاج قيمة كل منهما والتعويض بهذه القيمة في هذا

القانون ، كذلك فإن هذا القانون يستخدم في جميع أنواع التوزيعات التكرارية أي سواء كانت توزيعات مفتوحة أو توزيعات مقفولة .

مثال : من الجدول التكراري التالي أوجد معامل الالتواء :

الفئات	٠ -	٢٠ -	٤٠ -	٧٠ -	١٠٠ -	١٥٠ -	٢٠٠ -	٢٦٠ -	٣٢٠-٣٨٠
التكرارات	١٣	٢٧	٣٣	٤١	٥٣	٤٢	٣٦	٢٤	١٥

الحـــــــل

ف	ك	حدود عليا للفئات	تكرار متجمع صاعد
٠ -	١٣	أقل من ٢٠	١٣
٢٠-	٢٧	أقل من ٤٠	٤٠
٤٠-	٣٣	أقل من ٧٠	٧٣
٧٠	٤١	أقل من ١٠٠	١١٤
١٠٠-	٥٣	أقل من ١٥٠	١٦٧
١٥-	٤٣	أقل من ٢٠٠	٢١٠
٢٠٠-	٣٦	أقل من ٢٦٠	٢٤٦
٣٦٠-	٢٤	أقل من ٣٢٠	٢٧٠
٣٢٠-٣٨٠	١٥	أقل من ٣٨٠	٢٨٥

ترتيب الربيع الأدنى = (مجـ ك) / ٤ = ٢٨٥ / ٤ = ٧١,٢٥

قيمة الربيع الأدنى = ٤٠ + (٧١,٢٥ − ٤٠) / (٧٣ − ٤٠) × ٣٠

= ٤٠ + ٣١,٢٥ / ٣٣ × ٣٠

= ٤٠ + ٢٨,٤٠٩ = ٦٨,٤٠٩

ترتيب الربيع الأعلى = (مجـ ك) / ٤ × ٣ = ٢٨٥ / ٤ × ٣ = ٢١٣,٧٥

قيمة الربيع الأعلى = ٢٠٠ + (٣١٢,٧٥ − ٢١٠) / (٢٤٦ − ٢١٠) × ٦٠

= ٢٠٠ + ٣,٧٥ / ٣٦ × ٦٠

= ٢٠٠ + ٢٢٥ / ٣٦ = ٢٠٠ + ٦,٢٥

= ٢٠٦,٢٥

ترتيب الوسيط = (مجـ ك) / ٢ = ٢٨٥ / ٢ = ١٤٢,٥

قيمة الوسيط = ١٠٠ + (١٤٢,٥ − ١٤٤) / (١٦٧ − ١١٤) × ٥٠

$$= ١٠٠ + \frac{٢٨,٥}{٥٣} \times ٥٠$$

$$= ١٠٠ + \frac{١٤٢٥}{٥٣}$$

$$= ١٠٠ + ٢٦,٨٨٦ = ١٢٦,٨٨٦$$

معامل الالتـــــواء =

$$\frac{\text{(قيمة الربيع الأعلى – الوسيط) – (الوسيط – قيمة الربيع الأدنى)}}{\text{(قيمة الربيع الأعلى – الوسيط) + (الوسيط – قيمة الربيع الأدنى)}}$$

$$= \frac{(٢٠٦,٢٥٠ - ١٢٦,٨٨٦) - (١٢٦,٨٨٦ - ٦٨,٤٠٩)}{(٢٥٠,٢٠٦ - ١٢٦,٨٨٦) + (١٢٦,٨٨٦ - ٦٨,٤٠٩)}$$

$$= \frac{٨٥,٤٧٧ - ٧٩,٣٦٤}{٨٥,٤٧٧ + ٧٩,٣٦٤} = \frac{٢٠,٨٨٦}{١٣٧,٨٤١} = ٠,١٥١٥$$

الفصل الثامن

الانحـــدار Regression

الانحـــدار الخطي البســـيط :

يتناول تحليل الانحدار البسيط دراسة العلاقة بين متغيرين أحدهما هـــو المتغير التابع Dependent Variable وهو المتغير الذي يهتم الباحث بدراسته أصلاً ويحاول التنبؤ به وذلك من خلال دراسة علاقته بالمتغير الآخـــر ، أمـــا المتغير الثاني ويسمى المتغير المستقل وهو يهتم بالتغيرات الحادثة في المتغير التابع ، فعلى سبيل المثال عند دراسة علاقة الدخل بالاستهلاك ، فـــإن الـــدخل يمثل المتغير المستقل والاستهلاك يمثل المتغير التابع ، فإذا ما توافرت بيانـــات عن دخل عينة من المستهلكين ومتوسط الاستهلاك لكل منهم فإنه يمكن تقـــدير تلك العلاقة الانحدارية والاعتماد عليها في التنبؤ بالاستهلاك المتوقع .

ويمكن كتابة هذه العلاقة في الصور التالية :

$$Y = f(X)$$

أي أن (Y) دالة للمتغير (X)

والغرض من تحليل الانحدار الخطي البسيط هو الوصول إلى معادلة رياضـــية خطية ، يمكن من خلالها تحديد الخط البياني الذي يصف العلاقة المتوسطة بين المتغيرين أو تقدير معادلة انحدار (Y) على (X) وتكتب على النحو التالي :

$$(Y/X)$$

وتأخذ معادلة الخط المستقيم الصورة الرياضية التالية :

$$\hat{Y} i = \hat{a} + \hat{B} x i$$

حيث :

القيمة التقديرية للمتغير التابع في المشاهدة (i) = $\hat{Y} i$

قيمة المتغير المستقل في المشاهدة (i) = $X i$

مقطع ، أو قيمة المتغير (Y) عندما تكون قيمة (X) مساوية للصفر = $\hat{a}$

انحدار الخط المستقيم ، أو مقدار التغير الذي يحدث في العامل : $\hat{B}$

التابع (Y) نتيجة تغير العامل المستقل (X) بوحدة واحدة .

ويمكن تقدير ثوابت المعادلة الخطية باستخدام طريقة المربعات الصغرى .

طريقة المربعات الصغرى :

إن إحداثيات النقط التي يتم توقيعها على الرسم لا تقع جميعها على الخط الذي نوفقه وعلى ذلك فتكون هناك بعض النقط التي تنحرف عن هذا الخط وتستخدم طريقة المربعات الصغرى لتوفيق خط مستقيم (أو منحنى) وبحيث أن يكون مجموع مربعات انحرافات النقط عن هذا الخط أصغر ما يمكن .

ولفهم طريقة المربعات الصغرى ، نفرض أننا لدينا مجموعة من القيم المشاهدة للمتغيرين س ، ص وكانت هذه القيم هي (س ، ص) ، (س^{٢} ، ص^{٢}) ، (س^{٣} ، ص^{٣}) ، ...... (س$_{ن}$ ، ص$_{ن}$) والمطلوب هو توفيق خط مستقيم لمجموعة هذه القيمة كما هو مبين بالشكل (١) .

لذلك نفرض أن أحسن توفيق للخط المستقيم المطلوب هو الممثل بالمعادلة التالية :

ص = م س + جـ (١)

حيث :

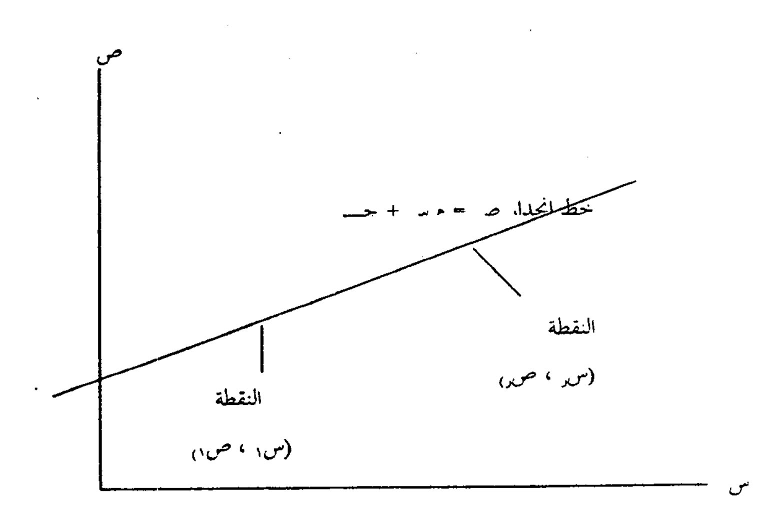
ص : متغير تابع

س : متغير مستقل

م : ميل المستقيم على المحور الأفقي (معامل انحدار ص على س)

جـ : كمية ثابتة (مقدار ثابت) هو طول الجزء الذي يقطعه المستقيم من المحور الرأسي .

وتسمى المعادلة السابقة بمعادلة انحدار ص على س أو تكتب ص/س ويكون م ، ج مجهولان في هذه المعادلة وتحديد قيمتها هو الذي يجعل مجموع مربعات انحرافات النقطة عن الخط المستقيم (أ) أصفر ما يمكن .



واستخدام طريقة المربعات الصغرى يوصلنا إلى الحصول على المعادلتين الطبيعيتين الآتيتين :

مجـــ ص = م مجـــ + ن ج .... (٢)

مجـــ س ص = م جـــ س^{٢} + جـــ مجـــ س .... (٣)

ولما كانت جميع قيم س ، ص معلومة وهي القيم المشاهدة وكـــذلك ن (عدد القيم) معلومة فإنّ جميع القيم في المعادلتين (٢) ، (٣) تكون معلومة مـــا عدا قيمة م ، جـــ وهاتين القيمتين يمكن الحصول عليهما عـــن طريـــق حـــل المعادلتين آنياً وبالتعويض عنهما في معادلة المستقيم = ص م س + ج نحصل على المعادلة المطلوبة .

أولاً : إيجاد معادلة الانحدار من بيانات غير مبوبة :

مثال (١) :

البيانات التالية تمثل الكميات المطلوبة من سلعة معينة (س) والسعر المناظر لها (ص) :

الكميات المطلوبة (س)	٣	٥	٧	٧	٨
السعر (ص)	٤	٦	٦	١٥	٩

والمطلوب إيجاد معادلة انحدار السعر على الطلب أي معادلة انحـــدار ص/س

الحـــل :

لإمكان حــل المعادلتين (٢) ، (٣) فيلزم إيجــاد قيـــم مجـــ س ، مجــ ص ، مجــ $س^٢$ ، مجــ س ص والجدول التالي يوضح ذلك .

س	ص	$س^٢$	س ص	$ص^٢$
٣	٤	٩	١٢	١٦
٥	٦	٢٥	٣٠	٣٦
٧	٦	٤٩	٤٢	٣٦
٧	١٥	٤٩	١٠٥	٢٢٥
٨	٩	٦٤	٧٢	٨١
٣٠	٤٠	١٩٦	٢٦١	٣٩٤

والآن بتعويض القيم في المعادلتين (١) ، (٣) نحصل على :

٤٠ = ٣٠ م + ٥ جــ

٢٦١ = ١٩٦م + ٣٠ جــ

ولحل هاتين المعادلتين نضرب الأولى في ٦ ونطرحها مــن المعادلــة الثانية فنحصل على :

٢٦١ = ١٩٦م + ٣٠ جــ

(-) ٢٤٠ = ١٨٠م + ٣٠ جــ

٢١ = ١٦م

ومنها م = $\frac{٢١}{١٦}$ = ١,٣١

وللحصول على قيمة جـ نعوض بقيمة م في أي من المعادلتين فإذا عوضنا في المعادلة الأولى نحصل على :

٤٠ = ٣٠ × ١,٣١ + ٥ جـ

٤٠ = ٣٩,٣ + ٥جـ وهذا يعني أن

٥ جـ = ٤٠ – ٣٩,٣ = ٠,٧ وبذلك تكون :

$$جـ = \frac{٠,٧}{٥} = ٠,١٤$$

فتكون معادلة انحدار ص / س هي :

ص = ١,٣١س + ٠,١٤

هذا ويمكن استخدام وسط فرضي للمتغير س ووسط فرضي آخر للمتغير ص دون أن يؤثر ذلك على النتيجة .

طريقة أخرى لحساب م ، جـ :

يمكن إيجاد معامل الانحدار (م) جبرياً من المعادلتين الطبيعيتين (٢) ، (٣) وذلك بضرب طرفي المعادلة (٢) في مجـ س والمعادلة (٣) في ن ينتج أن :

مجـ س مجـ ص = م (مجـ س)٢ + ن جـ مجـ س ......(٤)

ن مجـ س ص = ن م مجـ س ٢ + ن جـ مجـ س ......(٥)

بطرح المعادلة (٤) من المعادلة (٥) ينتج أن :

ن مجـ س ص – مجـ س مجـ ص = ن م مجـ س٢ – م (مجـ س)٢

= م (ن مجـ س ٢ – (مجـ س)٢)

وهذا يعني أن

$$م = \frac{\frac{1}{ن}(\text{مجـ } س\ ص) - س\ ص}{ع^{2}_{س}}$$

وبقسمة حدود كل من البسط والمقام على ن ينتج أن :

$$م = \frac{\frac{1}{ن}(\text{مجـ } س\ ص) - \bar{س}\ \bar{ص}}{ع^{2}_{س}} \quad ....(٦)$$

وهذه الصيغة الأخيرة تعطينا مباشرة قيمة (م) معامل انحـدار ص/س فمن البيانات للمثال السابق نجد أن :

$$م = \frac{\frac{٢٦١}{٥} - ٦ \times ٨}{\frac{١٩٦}{٥} - (٦)^{٢}} = \frac{٥٢,٢ - ٤٨}{٣٩,٢ - ٣٦}$$

$$= \frac{٤,٢}{٣,٢} = ١,٣١$$

وهي نفس الإجابة السابقة .

ويمكن إذن استنتاج قيمة جـ فبقسمة المعادلة (٢) على ن ينتج أن :

$$\frac{\text{مجـ } ص}{ن} = م \frac{\text{مجـ } س}{ن} + جـ$$

ص = م س + جـ

وهذا يعني أن :

جـ = $\bar{ص}$ − م $\bar{س}$	..........(٧)

ولما كان $\bar{ص}$ = $\frac{٤٠}{٥}$ = ٨ ، $\bar{س}$ = $\frac{٣٠}{٥}$ = ٦ ، م = ١,٣١

وبالتعويض في المعادلة (٧) نجد أن :

جـ = ٨ − (١,٣١) × ٦

= ٨ − ٧,٨٦

= ٠,١٤

معادلة انحدار س على ص :

في المعادلة (١) اعتبرنا ص متغير تابع و س متغير مستقل وفي هذه الحالة يعطى خط الانحدار ص/س أحسن قيم ص بدلالة س عندما تكون س مضبوطة وقيم ص عرضة للخطأ ، أما النوع الثاني الذي سنتناوله الآن هو لو اعتبرنا س متغير تابع وص هي المتغير المستقل فيكون خط الانحدار الناشئ عن ذلك هو خط انحدار س/ص وهذا الذي يعطي أحسن قيم س بدلالة ص عندما تكون ص مضبوطة وقيم س عرضة للخطأ .

ومعادلة خط انحدار س على ص يمكن أن تكتب على الصورة التالية :

س = أ ص + ب	.... (٨)

وبإتباع طريقة المربعات الصغيرة فتكون المعادلتان الطبيعيتان هما :

$$\text{مجـــ س} = \text{أ مجـــ ص} + \text{ن ب}$$

$$\text{مجـــ س ص} = \text{أ مجـــ ص}^{٢} + \text{ب مجـــ ص}$$

ومن هاتين المعادلتين يمكن الوصول إلى قيمة أ حيث :

$$\text{أ} = \frac{\frac{١}{\text{ن}}\ \text{مجـــ س ص} - \bar{\text{س}}\ \bar{\text{ص}}}{\text{ع}^{٢}_{\text{س}}} \quad ....(٩)$$

$$\text{ب} = \bar{\text{س}} - \text{أ}\ \bar{\text{ص}} \quad ....(١٠)$$

مثال (٢) :

في المثال السابق (مثال ١) ، أوجد معادلة خط انحدار س/ص .

الحـــــل :

بالرجوع إلى الجدول السابق وباستخدام العلاقة (٩) نجد أن :

$$١ = \frac{\frac{٢٦١}{٥} - ٦ \times ٨}{\frac{٣٩٤}{٥} - (٨)^{٢}} = \frac{٥٢,٢ - ٤٨}{٧٨,٨ - ٦٤} = \frac{٤,٢}{١٤,٨} = ٠,٢٨$$

وحيث أن $\bar{\text{س}}$ = ٦، $\bar{\text{ص}}$ = ٨ فمن المعادلة (١٠) نجد أن :

$$\text{ب} = ٦ - (٠,٢٨ \times ٨)$$

$$= ٣,٧٦$$

وبالتعويض بقيمتي أ ، ب في معادلة س/ص (معادلة ٨) نحصل على :

س = ٠,٢٨ ص + ٣,٧٦

فائدة خط الانحـــدار :

يتضح ممّا سبق أن خط الانحدار يترجم العلاقــة بـــين المتغيـــرين أو الظاهرتين في صورة رياضية . وفي الحقيقة أن الغرض من الحصول علــى هذه العلاقة الخطية هو التنبؤ بقيمة أحد الظاهرتين إذا علمت قيمـــة الظـــاهرة الأخرى . ففي معادلة انحدار ص/س على الصورة ص = م س + جـ فللتنبؤ بقيمة ص إذا علمت قيمة س فيتم تحديد قيمتي م ، جـ كما سبق أن ذكرنا ثـــم نحصل على القيمة المتنبأ بها للمتغير ص بمعلومية قيمة س ، وبالمثـــل ففـــي معادلة انحدار س/ص على الصورة : س = أ ص + ب فللتنبؤ بقيمـــة س إذا علمت قيمة ص فيتم تحديد قيمتي أ ، ب ثم نعوض في المعادلـــة بعـــد ذلـــك للحصول على القيمة المتنبأ بها للمتغير س بمعلومية قيمة ص .

كما يلاحظ أن حاصل ضرب معاملي انحدار ص على س ، س علـــى ص يساوي مربع معامل الارتباط وأن المعاملات الثلاثة متفقة في الإشارة .

أي أنه بمعرفة معامل الارتباط والانحراف المعياري لقيم س ، ولقـــيم ص وكذلك الوسط الحسابي لقيم س ، ص يمكن إيجاد معادلتي خط الانحـــدار مباشرة .

مثال (٣) :

من نتائج بيانات المثالين السابقين (١ ، ٢) أوجـــد :

- معامل الارتباط وتحقق من صحة النتائج عن طريق معامل الانحدار.

الحــــل :

نعلم أن معامل الارتباط (ر) يحسب عن طريق الصيغة التالية (قـــانون بيرسون)

$$\text{ر} = \frac{\frac{1}{\text{ن}}\,\text{مجـ س ص} - \bar{\text{س}}\,\bar{\text{ص}}}{\sqrt{\text{ع}^2_{\text{س}} \times \text{ع}^2_{\text{ص}}}}$$

بالرجوع إلى البيانات في الجدول الأخير نحصل على :

$$\text{ر} = \frac{\frac{\text{٢٦١}}{\text{٥}} - \text{٦} \times \text{٨}}{\sqrt{\left(\frac{\text{١٩٦}}{\text{٥}} - (\text{٦})^2\right)\left(\frac{\text{٣٩٤}}{\text{٥}} - (\text{٨})^2\right)}}$$

$$= \frac{\text{٤,٢}}{\sqrt{\text{٣,٢} \times \text{١٤,٨}}} = \frac{\text{٤,٢}}{\text{٦,٨٨٢}} = \text{٠,٦١}$$

ومن ناحية أخرى فبالرجوع إلى قيمتي أ ، م من المثالين السابقين نجد أن:

$\text{ر}^2 = \text{أ م}$

$= \text{٠,٢٨} \times \text{١,٣٣}$

= ٠,٣٦٦٨

إذاً ر = $\sqrt{٠,٣٦٦٨}$ = - ٠,٦١

وأخذنا الإشارة السالبة من هذا الجذر لأن معاملي الانحدار كانت لهما نفس الإشارة السالبة .

ثانياً : إيجاد معامل الانحدار من بيانات مبوبة :

يمكن إيجاد معادلة انحدار ص/س أو س/ص بدلالة معامل الارتباط الذي تم إيجاده من بيانات مبوبة سابقاً .

إيجاد معادلة انحدار ص / س بدلالة معامل الارتباط لبيانات مبوبة :

أوضحنا سابقاً أن معادلة خط انحدار ص/س تأخذ الشكل الآتي :

ص = م س + جـ

يمكن إيجاد (م) وهي ميل المستقيم على المحور الأفقي وهو معامل انحدار ص/س بدلالة معامل الارتباط الذي تم إيجاده سواء إذا كانت البيانات مبوبة أو غير مبوبة باستخدام العلاقة الآتية :

$$م_{ص/س} = ر \times \frac{ع_{ص}}{ع_{س}}$$

حيث ر هي معامل الارتباط بين المتغيرين .

$ع_{س}$ ، $ع_{ص}$ هما الانحراف المعياري لكل من المتغيرين س ، ص ويمكن إيجادهما من البيانات المبوبة .

ويمكن إيجاد جـ وهو طول الجزء الذي يقطعه المستقيم من المحور الرأسي من العلاقة التالية :

$$جـ = \overline{ص} - م \overline{س}$$

حيث س ، ص هما الوسط الحسابي لبيانات الظاهرة س ، ص .

الفصل التاسع

الارتباط CORRELATION

لدراسة الارتباط بين متغيرين نحتاج إلى مقياس يقيس لنا درجة العلاقة بينهما واتجاه هذه العلاقة بمعنى أنه إذا وجدنا الزيادة في أحد المتغيرين يصاحبها زيادة في المتغير الثاني كان هذا دليلاً على وجود علاقة بينهما والعلاقة في هذه الحالة تسمى علاقة طردية بما يعني أن المتغيرين يسيران في اتجاه واحد فزيادة أحدهما يصحبها زيادة في المتغير الثاني ونقص أحدهما يصاحبها نقص في المتغير الثاني أما إذا كانت الزيادة في أحدهما يصاحبه نقص في المتغير الثاني فهذا دليلاً أيضاً على وجود علاقة بينهم ولكن العلاقة في هذه الحالة تسمى علاقة عكسية بمعنى أن زيادة أحدهم يصاحبه نقص في المتغير الثاني ونقص أحدهما يصاحبه زيادة في المتغير الثاني والأمثلة على ذلك كثيرة ومتعددة ، فلاشك من وجود علاقة طردية بين عدد الكيلومترات التي تقطعها السيارة وكمية الوقود المستخدمة ولاشك من وجود علاقة بين عدد ساعات الاستذكار والدرجة المتحصلة في الاختبار كذلك هناك علاقة بين طول الطفل وعمره وكل هذه العلاقات علاقات طردية ، أما العلاقات العكسية ، فمنها العلاقة بين الكمية المعروضة من سلعة معينة وسعرها فمن المعروف اقتصادياً أنه كلما زاد المعروض من سلعة معينة كلما أنخفض سعرها وأيضاً العلاقة بين عمر إطار السيارة وعدد الكيلو مترات المقطوعة يومياً وهناك العلاقة بين عمر المصباح الكهربائي وعدد ساعات استعماله يومياً .... الخ .

معامل ارتباط بيرســـــون
(PEARSON CORRELATION COEFFICIENT)

ذكرنا أننا بحاجة إلى مقياس يقيس لنا درجة العلاقة بـــين المتغيـــرين واتجاهها . هذا المقياس هو ما يطلق عليه معامل ارتباط بيرسون (ر) ومعامل الارتباط ببساطة شديدة هو كسر يتراوح بين ١ ، -١ ويســـتحيل أن يأخـــذ أي قيمة تزيد عن الواحد الصحيح ويستحيل أيضاً أن يأخذ أي قيمة تقل عن سالب واحد وممكن أن يساوي + ١ في حالة الارتباط الطردي التام ويمكن أن يساوي - ١ في حالة الارتباط العكسي التام وكلما قربت قيمة هذا الكسر مـــن الواحـــد الصحيح الموجب أو السالب كلما كانت العلاقة بين المتغيـــرين الـــداخلين فـــي حسابه قوية وكلما قربت قيمة هذا الكسر من الصفر سواء كان سالباً أو موجباً كان ذلك دليل على ضعف العلاقة بين المتغيرين .

- ١ ≥ معامل الارتباط" ≥ ١

أي أن إشارة هذا الكسر (معامل الارتباط) تدلنا على اتجاه العلاقة بين المتغيرين هل هي طردية أم عكسية فالإشارة الموجبة تعني أن العلاقة طرديـــة والإشارة السالبة تعني أن العلاقة عكسية . وينتهي دور الإشارة عند هذا الحـــد فهي تدلنا على طردية العلاقة أو عكسيتها فقط .

ولكن معرفة قوة العلاقة يأتي لنا من معرفة قيمة هذا الكسر هل هو قريب مـــن الواحد الصحيح ٠,٩ أو ٠,٨ مثلاً فنستدل أن العلاقة قوية بصرف النظر عـــن كونها طردية أو عكسية أم هل هو أي الكسر قريب من الصـــفر ٠,٢ أو ٠,١ مثلاً فنستدل أن العلاقة ضعيفة . ولإيجاد معامل الارتباط بين متغيـــرين س ،

ص مثلاً لابد وأن يكون لدينا أزواج من القيمة كل زوج من هذه القيم يمثل قيمة لـــ س و أخرى لـــ ص . فنضع هذه الأزواج من القيم فـــي جـــدول ونطبـــق القانون التالي :ـــ

$$\text{ر} = \frac{\text{ن مجـ س ص} - \text{مجـ س مجـ ص}}{\sqrt{\left[\text{ن مجـ س}^{٢} - (\text{مجـ س})^{٢}\right]\left[\text{ن مجـ ص}^{٢} - (\text{مجـ ص})^{٢}\right]}}$$

حيث ر هو معامل الارتباط .

مجـــ س مجموع قيم المتغير الأول س .

مجـــ ص مجموع قيم المتغير الثاني ص .

مجـــ $\text{س}^{٢}$ مجموع مربعات قيم المتغير الأول .

مجـــ $\text{ص}^{٢}$ مجموع مربعات قيم المتغير الثاني .

$(\text{مجـ س})^{٢}$ مربع مجموع قيم المتغير الأول .

$(\text{مجـ ص})^{٢}$ مربع مجموع قيم المتغير الثاني .

ن عدد أزواج القيم

والطريقة سهلة للغاية كما سيتضح من المثال التالي مع ملاحظة أن هناك فارق كبير بين مجـــ $\text{س}^{٢}$ ، $(\text{مجـ س})^{٢}$.

مثال (١) :

أختير عشوائياً عشرة طلاب ورصدت درجاتهم في اختبارين لمـــادتي الكيمياء والأحياء وكانت درجاتهم كالتالي والمطلوب معرفة هل هناك علاقـــة بين المادتين بمعنى أن تفوق الطالب في أحد المادتين يصحبه تفوق في المـــادة

الأخرى أم يصحبه تدني لمستواه في المادة الثانية علماً بأن اختبار الكيمياء كان من ١٠٠ واختبار الأحياء كان من ٥٠

اسم الطالب	محمد	عادل	رضا	اسامه	وائل	عمر	جابر	نبيل	خلف	عيسى
درجة الكيمياء	٦٩	٧١	٦٨	٦٤	٦٧	٦٦	٧٠	٦٢	٦٨	٦٥
درجة الأحياء	٢٨	٣٠	٣١	٢٥	٢٨	٢٥	٢٨	٢٦	٢٩	٢٨

الحـــــل :

للإجابة على السؤال في هذا المثال نحسب معامل الارتبـــاط ونرمـــز لدرجة الكيمياء بالرمز (س) ولدرجة الأحياء بالرمز (ص) ونكون جدول مـــن خمسة أعمدة للحصول على مجـــ س ، مجـــ ص ، مجـــ $س^2$ ، مجـــ $ص^2$ ، مجـــ س ص مع ملاحظة أنه لا علاقة للدرجة النهائية سواء في مادة الكيميـــاء أو في مادة الأحياء بحساباتنا لأنه كما سنرى يمكن اختصار هذه الأرقام ولكـــن نبدأ الآن بالأرقام الأصلية كالتالي :

س	ص	$س^{٢}$	$ص^{٢}$	س ص
٦٩	٢٨	٤٧٦١	٧٨٤	١٩٣٢
٧١	٣٠	٥٠٤١	٩٠٠	٢١٣٠
٦٨	٣١	٤٦٢٤	٩٦١	٢١٠٨
٦٤	٢٥	٤٠٩٦	٦٢٥	١٦٠٠
٦٧	٢٨	٤٤٨٩	٧٨٤	١٨٧٦
٦٦	٢٥	٤٣٥٦	٦٢٥	١٦٥٠
٧٠	٢٨	٤٩٠٠	٧٨٤	١٩٦٠
٦٢	٢٦	٣٨٤٤	٦٧٦	١٦١٢
٦٨	٢٩	٤٦٢٤	٨٤١	١٩٧٢
٦٥	٢٨	٤٢٢٥	٧٨٤	١٨٢٠
٦٧٠	٢٧٨	٤٤٩٦٠	٧٧٦٤	١٨٦٦٠

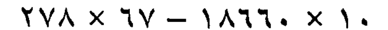
حصلنا من هذا الجدول على المجاميع المطلوبة لحساب معامل الارتباط

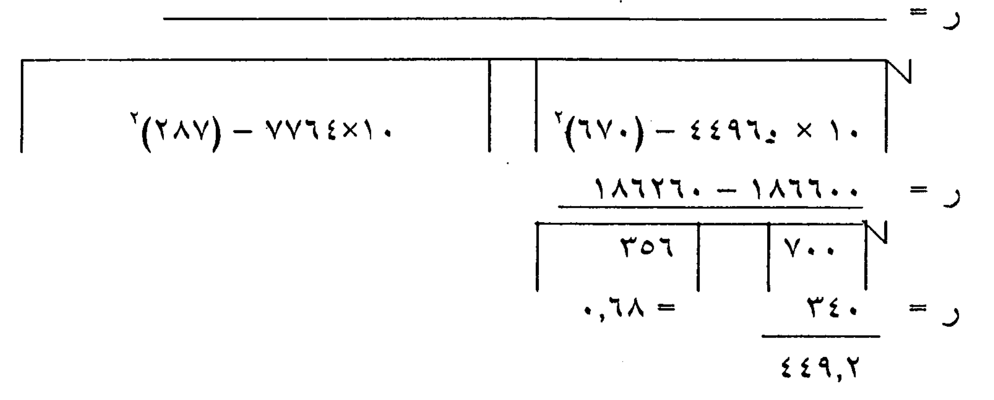
مجـ س = ٦٧٠ ، مجـ ص = ٢٧٨ ، مجـ س ص = ١٨٦٦٠

مجـ س٢ = ٤٤٩٦٠ ، مجـ ص٢ = ٧٧٦٤

ونطبق القانون :

$$ر = \frac{\text{ن مجـ س ص} - \text{مجـ س مجـ ص}}{\sqrt{\text{ن مجـ س}^{٢} - (\text{مجـ س})^{٢}} \, \sqrt{\text{ن مجـ ص}^{٢} - (\text{مجـ ص})^{٢}}}$$





$$ر = \frac{١٠ \times ١٨٦٦٠ - ٦٧ \times ٢٧٨}{\sqrt{١٠ \times ٤٤٩٦٠ - (٦٧٠)^٢} \quad \sqrt{١٠ \times ٧٧٦٤ - (٢٨٧)^٢}}$$

$$ر = \frac{١٨٦٦٠٠ - ١٨٦٢٦٠}{\sqrt{٧٠٠} \quad \sqrt{٣٥٦}}$$

$$ر = \frac{٣٤٠}{٤٤٩,٢} = ٠,٦٨$$

حيث أن معامل الارتباط موجب فإننا نستطيع أن نقول أن هناك علاقة طردية بين مادتي الكيمياء والأحياء وحيث أن قيمته ٠,٦٨ فإن المسافة بينه وبين الواحد الصحيح أقل بكثير من المسافة بينه وبين الصفر لذا فهو معامل ارتباط قوي بمعنى أن هناك علاقة قوية بين مادتي الكيمياء والأحياء ولكنها ليست قوية جداً ويمكن وصف العلاقة بأنها قوية جداً إذا كان معامل الارتباط ٠,٩ حتى ٠,٨

وهذا يجرنا إلى وضع مؤشرات لتفسير معامل الارتباط كالتالي :

معامل الارتباط	وصف العلاقة
١	ارتباط تام
٠,٩ – ٠,٨	ارتباط قوي جداً
٠,٧ – ٠,٦	ارتباط قوي
٠,٥	ارتباط متوسط

٠,٣ – ٠,٤	**ارتباط ضعيف**
٠,١ و ٠,٢	**ارتباط ضعيف جداً**
صفر	**لا يوجد ارتباط خطي**
سالب	**الارتباط عكسي**
موجب	**الارتباط طردي**

في مثال (١) السابق كانت العمليات الحسابية كبيرة للغاية وبدون الاستعانة بالآلة الحاسبة كان من الصعب الحصول على قيمة معامل الارتباط ولكن لمعامل الارتباط خاصية هامة جداً هو أنه لا يتأثر إذا طرحنا أي قيمة ثابتة من جميع قيم س وطرحنا أي قيمة ثابتة أخرى من جميع قيم ص ونستعمل القيم الجديدة بعد الطرح على أنها قيم س ، ص تماماً فلن يختلف معامل الارتباط والدليل على ذلك أننا لو طرحنا من جميع قيم س في الجدول السابق ٦٨ وطرحنا من جميع قيم ص ٢٨ – لاحظ أننا اخترنا ٦٨ ، ٢٨ لأنهما يتوسطان قيم س ، ص تقريباً لذلك فستكون الانحرافات أقل ما يمكن .

ويمكن إعادة حل المثال السابق كما يلي :

س	ص	$س^٢$	$ص^٢$	س ص
١	صفر	١	صفر	صفر
٢	٢	٩	٤	٦
صفر	٣	صفر	٩	صفر
-٤	-٣	١٦	٩	١٢
-١	صفر	١	صفر	صفر
٢	-٣	٤	٩	٦
٢	صفر	٤	صفر	صفر
-٦	-٢	٣٦	٤	١٢
صفر	١	صفر	١	صفر
-٣	صفر	٩	صفر	صفر
-١٠	-٢	٨٠	٣٦	٣٦

بعد الطرح يكون مجـــ س = -١٠ و مجـــ ص = -٢

مجـــ س٢ = ٨٠ ، مجـــ ص ٢ = ٣٦ ، مجـــ س ص = ٣٦

ونطبق القانـــــون

$$ر = \frac{١٠ \times ٣٦ - (-١٠)\ (-٢)}{\sqrt{\left[١ \times ٨٠ - (١٠)^٢\right]\left[١٠ \times ٣٦ - (٢)^٢\right]}}$$

$$ر = \frac{٣٦٠ - ٢٠}{\sqrt{٧٠٠ \times ٣٥٦}}$$

$$= \frac{٣٤٠}{٤٩٩,٢} = ٠,٦٨$$

وهي نفس النتيجة التي حصلنا عليها عند التعامل مع الأرقام الأصلية بدون إختزال وهذا يوضح ما ذكرناه عند بداية حل هذا المثال بأنه لا يعنينا في قليل أو كثير أن كانت درجات الكيمياء من ١٠٠ أو من غير ذلك لأننا ببساطة شديدة يمكن طرح أي قيم من جميع قيم المتغير س (الكيمياء) كما يمكن طرح أي قيمة من قيمة المتغير ص وهذه القاعدة عامة ويستفاد منها كثيراً لأنه يمكننا حساب معامل الارتباط بين الدخل وهو إما بالجنيه أو الدولار أو الريال أو أي عملة أخرى وبين عدد سنوات الخبرة وهو يقاس بالسنة ويمكن أيضاً عمل معامل الارتباط بين طول الشخص (سم) ووزنه (كجم) لأن أهم ما يعنينا هنا هو أزواج القيم أما مسألة وحدات قياس كل متغير فليس لها تأثير في حساب معامل الارتباط.

مثال (٢) :

إحسب معامل الارتباط بين الكمية المعروضة من التفاح وسعره في أحد الأسواق إذا علمت البيانات التالية :

الكمية المعروضة	السعر
١٥	٢٤
١٧	٢٢
٣٤	١٨
٨٠	١٥
٨٥	١٤
٩٢	١٢

الحـــل :

مجـ س = ٣٢٣ مجـ ص = ١٠٥

مجـ س ص = ٤٨٤٠

مجـ س^{2} = ٢٣٧٥٩ مجـ ص^{2} = ١٩٤٩

$$\text{ر} = \frac{٦ \times ٤٨٤٠ - ٣٢٣ \times ١٠٥}{\sqrt{\left\{ ٦ \times ٢٣٧٥٩ - \left\{٣٢٣\right\}^{2} \right\}\left\{ ٦ \times ١٩٤٩ - \left\{١٠٥\right\}^{2} \right\}}}$$

$$= \frac{٢٩٠٤٠ - ٣٣٩١٥}{٥٠٥٧}$$

$$= \frac{- ٤٨٧٥}{٥٠٥٧} = - ٠,٩٦$$

تفسير معامل الارتباط

حيث أن قيمة معامل الارتباط – ٠,٩٦ إذن العلاقة بين المتغيرين علاقة قويـــة جداً وكذلك نظراً لأن معامل الإرتباط سالب إذن العلاقة بين المتغيرين عكسية.

<u>معامل ارتباط الرتب (سبيرمان)</u>

أحياناً يكون المتغيران المراد قياس معامل الارتبـــاط بينهمـــا مقاســـان بمقياس ترتيبي بمعنى أن هذه المتغيرات متغيرات نوعية لا يمكن قياسها كميـــاً ولكنها تقاس كيفياً مثل تقديرات الطلاب ومثل أراء الناس تجاه مشـــكلة معينـــة ولحساب معامل الارتباط بينهما نستخدم معامل ارتباط آخر غير معامل بيرسون يطلق عليه معامل ارتباط الرتب (سبيرمان) وفي الواقع فإن معامـــل ارتبـــاط

الرتب (سبيرمان) هو صورة رياضية مشتقه من نفس الصورة الرياضية لمعامل ارتباط بيرسون ولكن ليس هنا مجال إثبات الصورة الرياضية وسنكتفي بكتابة قانون معامل ارتباط الرتب وشرح كيفية تطبيقه ، في مثال الطلاب العشرة السابق (مثال ١) لو كنا نريد حساب معامل الارتباط بين مادتي الطبيعة والرياضيات وكانت البيانات المتاحة لدينا هي تقديرات هؤلاء الطلاب في مادتي الرياضيات والطبيعة وليست درجاتهم فلمعرفة العلاقة بين هاتين المادتين فإننا نحول التقديرات إلى ترتيب ونحسب الفرق بين ترتيب كل طالب في كلا المادتين ونربع هذا الفرق لنحصل على ما نسميه مجـ ف٢ وهو المقدار الوحيد المطلوب حسابه لإيجاد معامل ارتباط الرتب حيث أن القانون في هذه الحالة هو:

$$ر = ١ - \frac{٦ \text{ مجـ } ف^{٢}}{ن (ن^{٢} - ١)}$$

حيث ن عدد المفردات (عدد أزواج القيم) .

ولنفرض أن البيانات كانت كالتالي :

مثال (٣)

اسم الطالب	محمد	عادل	رضا	اسامه	وائل	عمر	جابر	نبيل	خلف	عيسى
تقدير الرياضة	ض ج	ل	م	ل	ض	ج ج	ج	ض	ل	م
تقدير الطبيعة	ل	ج	ج ج	ل	ج	ل	م	ض ج	ض	ج ج

حيث : ممتاز (م) ، جيد جداً (ج ج) ، جيد (ج) ، مقبول (ل) ، ضعيف (ض) ، ضعيف جداً (ض ج)

ولحساب معامل ارتباط الرتب من هذا الجدول نعطي كل تقدير رتبه ممتاز (١) وجيد جداً (٢) ، جيد (٣) وهكذا ، مع ملاحظة أنه إذا تكرر تقدير معين نحسب متوسط المراكز التي كانوا سيحتلونها بمعنى أنه في المثال السابق في تقديرات مادة الطبيعة حصل جابر على تقدير ممتاز فيأخذ الرتبة (١) وحصل رضا وعيسى على تقدير جيد جداً فلا نستطيع أعطاء أحدهما الرتبة (٢) والآخر (٢) مكرر وأيضاً لا نستطيع أعطاء أحدهما الرتبة (٢) والآخر الرتبة (٣) ولكن كلاً منهم يأخذ الرتبة ٢,٥ وهي متوسط الرتبة ٢ والرتبة ٣ ولو كان هناك ثلاثة حاصلين على نفس التقدير نجمع مواقعهم وتقسم على ٣ وكل منهم يأخذ الترتيب الناتج بعد القسمة ونجري الحل كالتالي :ـ

اسم الطالب	تقدير الرياضيات	تقدير الطبيعة	رتب تقدير الرياضيات	رتب تقدير الطبيعة	الفروق ف	ف٢
محمد	ضعيف جداً	مقبول	١٠	٧	٣	٩
عادل	مقبول	جيد	٥,٥	٤,٥	١	١
رضا	ممتاز	جيد جداً	١	٢,٥	١,٥	٢٥٥
أسامه	مقبول	مقبول	٥,٥	٧	١,٥	٢,٢٥
وائل	ضعيف	جيد	٨,٥	٤,٥	٤	١٦
عمر	جيد جداً	مقبول	٢	٧	٥	٢٥
جابر	جيد	ممتاز	٣	١	٢	٤
نبيل	ضعيف	ضعيف جداً	٨,٥	١٠	١,٥	٢,٢٥
خلف	مقبول	ضعيف	٥,٥	٩	٣,٥	١٢,٢٥
عيسى	مقبول	جيد جداً	٥,٥	٢,٥	٣	٩

لاحظ هنا أن الفروق تحسب بإهمال الإشارة لأننا سنربعها فلا داعـــي للإشارة وكل ما يهمنا هو مجموع مربعات الفروق مجـــ ف^{٢} ونجد أنه في هذا المثال يساوي ٨٣ وبالتالي :ـــ

$$ر = ١ - \frac{٦ \text{ مجـ ف}^{٢}}{ن (ن^{٢} - ١)}$$

$$ر = ١ - \frac{٦ \times ٨٣}{١٠ \times ٩٩} = ٠,٥$$

تفسير معامل الارتباط في هذه الحالة :

حيث أن معامل الارتباط موجب فإنه توجد علاقة طردية بين المادتين وحيث أن قيمة معامل الارتباط = ٠,٥ فالعلاقة علاقة متوســطة ، أي أن العلاقــة بــين المادتين ليست بالقوية وليست بالضعيفة ، وإنما هي علاقة طردية متوسطة .

<u>معامل الارتباط الجزئـــي :</u>

إذا كان لدينا ثلاثة متغيرات س ، ص ، ع مثلاً ونريد حساب معامل الارتبــاط بين متغيرين منهما مع تثبيت أثر المتغير الثالث عليهما فهذا يعرف بالارتبــاط الجزئي ، فلو رمزنا للمتغير الأول س بــالرمز (١) ، والمتغيــر الثــاني ص بالرمز (٢) والمتغير الثالث ع بالرمز (٣) فيكون :ـــ

$$ر_{٢١,٣} = \frac{ر_{٢١} - ر_{٣١} \, ر_{٣٢}}{\sqrt{(١ - ر_{٣١})^{٢} \, (١ - ر_{٣٢})^{٢}}}$$

وهذا يعطينا معامل الارتباط بين س ، ص مع تثبيت أثر ع عليهما مع ملاحظة أن :

$\text{ر}_{٣١}$ معناه معامل الارتباط بين المتغير الأول والثالث .

$\text{ر}_{٢١}$ معناه معامل الارتباط بين المتغير الأول والثاني .

$\text{ر}_{٢١.٣}$ معناه معامل الارتباط الجزئي بين المتغير الأول والمتغير الثاني مع تثبيت أثر المتغير الثالث وبالمثل .

$$\text{ر}_{١,٣٢} = \frac{\text{ر}_{٣٢} - \text{ر}_{٣١}\,\text{ر}_{٢١}}{\sqrt{(١ - \text{ر}_{٣١})^{٢}\,(١ - \text{ر}_{٣٢})^{٢}}}$$

الفصل العاشـــــر

السلاسل الزمنية والأرقام القياسية

أولاً : السلاسل الزمنية :

السلاسل الزمنية تعتبر تطبيق مباشر على الانحدار حيث يتوفر لـــدينا مجموعة من البيانات عن ظاهرة معينة وتطورها عبر حقبة من الزمن وهــذه الظاهرة قد تكون تطور إنتاج مصنع معين وقد تكون أعداد السكان وقد تكــون صادرات دولة أو وارداتها ... الخ ، وطبيعي أن المتغير المستقل هنا هو الزمن والمتغير التابع هو تطور الظاهرة والطريقة هنا تماماً كما فعلنـــا فـــي حالـــة الانحدار وهي تقدير معادلة خط مســـتقيم ص = أ + ب س مـــع اعتبـــار أن السنوات هي المتغير س والظاهرة هي المتغير ص .

ونستخدم نفس القوانين السابقة في الانحدار لتقدير قيمة الثوابت أ ، ب ولكننا دائماً نختصر أرقام السنوات بطرح سنة وسيطة من جميع أرقام السنوات للحصول على قيم صغيره للمتغير س كما سيتضح من المثال التالي .

مثال (١)

كيف يمكن تقدير إنتاج السعودية من البترول عـــام ١٤٢٠ إذا كـــان الإنتـــاج بالمليون برميل في السنوات السابقة كالتالي :

العـــام	الإنتـــاج
١٤٠٨	١٥
١٤٠٩	١٧
١٤١٠	٢٠
١٤١١	٢٣
١٤١٢	٢٦
١٤١٣	٣٠

الحـــل :

العام	الإنتاج	س	س ص	$س^٢$
١٤٠٨	١٥	– ٢	– ٣٠	٤
١٤٠٩	١٧	– ١	– ١٧	١
١٤١٠	٢٠	صفر	صفر	صفر
١٤١١	٢٣	١	٢٣	١
١٤١٢	٢٦	٢	٥٢	٤
١٤١٣	٣٠	٣	٩٠	٩
	١٣١	٣	١١٨	١٩

$$ب = \frac{ن\ مجـ\ س\ ص - مجـ\ س\ مجـ\ ص}{ن\ مجـ\ س^{٢}\ (مجـ\ س)^{٢}}$$

$$أ = \frac{مجـ\ ص}{ن} - ب\ \frac{مجـ\ س}{ن}$$

$$أ = \frac{١٣١}{٦} - ٣ \times \frac{٣}{٦} = ٢١,٨ - ١,٥ = ٢٠,٣$$

إذن ص = ٢٠,٣ + ٣ س

الإنتاج عام ١٤٢٠ = ٢٠,٣ + ٣ × ١٠ = ٢٠,٣ + ٣٠

= ٥٠,٣ مليون برميل

عوضنا هنا عن العام ١٤٢٠ بالقيمة ١٠ حيث أن عام ١٤١٣ أخذ القيمة ٣ فيكون عام ١٤١٤ = ٤ وعام ١٤١٥ = ٥ .... وهكذا حتى عام ١٤٢٠ = ١٠ وذلك لأننا طرحنا عام ١٤١٠ من جميع السنوات .

ثانياً : الأرقام القياسية **INDEXNUMBERS**

تهتم بقياس التغير في الظواهر المحيطة بنا وبالذات التغير في الظواهر الاقتصادية ، وتعتبر أسعار السلع المختلفة من أكثر الظواهر تغيراً في ميدان الظواهر الاقتصادية ، حيث في مجال الإنتاج نقارن دائماً بين كمية الإنتاج من منتج معين الآن وكمية نفس المنتج في العام الماضي . وهذا أمراً طبيعياً لأن الظواهر الاقتصادية عرضه للتغير من وقت لآخر ومن مكان إلى مكان آخر وأسعار السلع بوجه خاص عرضه لتغيرات كثيرة من عام لعام آخر أو نتيجة لتغيرات العرض والطلب على هذه السلع ونتيجة لتغيرات تكاليف الإنتاج والنقل والتخزين والتأمين ... الخ ولدراسة هذه التغيرات نحتاج إلى مقياس إحصائي

يبين لنا كمية هذا التغير ونوعيته من عام إلى آخر لنتمكن من مقارنة الظواهر الاقتصادية من فترة إلى فترة أخرى أو من مكان إلى مكان آخر . هذا المقياس هو ما يعرف بالرقم القياسي .

وعلى ذلك فالرقم القياسي يمكن أن يقيس لنا التغير الذي طرأ على أسعار سلع التصدير مثلاً من العام السابق إلى العام الحالي أو التغير الذي طرأ على أجور العمال في منطقة معينة من عام إلى عام أو التغير في أجورهم من مكان إلى مكان آخر في نفس العام والأمثلة على ذلك كثيرة ومتعددة ونسوق منها الآتي :

- **مثال (٢)**

كان إنتاج أحد مصانع الملابس الجاهزة ٣٠٠,٠٠٠ قطعة في عام ١٩٩٠ إرتفع هذا الإنتاج إلى ٤٥٠,٠٠٠ قطعة في عام ١٩٩٣ ، إحسب الرقم القياسي للإنتاج .

- **الحل :**

الرقم القياسي للإنتاج $= \frac{٤٥٠٠٠٠}{٣٠٠٠٠٠} \times ١٠٠ = ١٥٠\%$

وهذا يعني أنه إذا كان الإنتاج عام ١٩٩٠ = ١٠٠ وحده فإن الإنتاج عام ١٩٩٣ = ١٥٠ وحده .

أي أن الإنتاج زاد بنسبة ٥٠% .

- **مثال (٣)**

إذا كانت تكلفة إنتاج السيارة في مصانع هوندا للسيارات ٨٠٠٠ دولار في عام ١٩٩٠ وكانت تكلفة إنتاج نفس السيارة في عام ١٩٩٣ ١٠٠٠٠ دولار فما هو الرقم القياسي للتكلفة .

- **الحـــل :**

الرقم القياسي للتكلفة = $\frac{١٠٠٠٠}{٨٠٠٠}$ × ١٠٠ = ١٢٥% وهذا يعنـــي إرتفـــاع التكلفة بنسبة ٢٥% من عام ٩٠ إلى عام ٩٣ .

- **مثال (٤)**

إذا كان سعر جهاز الكمبيوتر في عام ١٩٩٠ ٤٠٠٠ جنيه إنخفــض إلـــى ٢٥٠٠ جنيه في عام ١٩٩٤ فما هو الرقم القياسي لسعر الكمبيوتر .

الحـــل :

الرقم القياسي = $\frac{٢٥٠٠}{٤٠٠٠}$ × ١٠٠ = ٦٢,٥%

وهذا يعني أنه إذا كان جهاز الكمبيوتر في عام ١٩٩٠ ثمنه ١٠٠ جنيه فإنه في عام ١٩٩٣ أصبح ثمنه ٦٢,٥ جنيه أي أن الســعر أنخفــض بنســبة ٣٧,٥% (١٠٠ – ٦٢,٥) .

من هذه الأمثلة يمكن أن نستنتج الآتـــي :

الرقم القياسي للمتغير = $\frac{\text{قيمة المتغير في سنة المقارنة}}{\text{قيمة المتغير في سنة الأساس}}$ × ١٠٠

وإذا كان الجواب أكبر من ١٠٠ فإن ذلك يعني أن قيمة هذا المتغير إرتفعت ما بين سنة الأساس وسنة المقارنة بنسبة تعادل هذا الجواب + ١٠٠ .
وإذا كان الجواب أقل من ١٠٠ فإن ذلك يعني أن قيمة هذا المتغير إنخفضت ما بين سنة الأساس وسنة المقارنة نسبة تعادل (١٠٠- الجواب).

الرقم القياسي لأسعار مجموعة من السلع

نفرض أن لدينا مجموعة من البيانات عن أسعار بعض الفواكه والكميات المستهلكة منها في أعوام ١٩٩٠ ، ١٩٩٣ .
سنعتبر سنة ٩٠ هي سنة الأساس وسنة ٩٣ هي سنة المقارنة وسنقوم بعرض لجميع الأرقام القياسية الشائعة الاستخدام باستعمال هذه المجموعة من البيانات الموضحة في الجدول التالي :

السلعة	الأسعار		الكميات	
	١٩٩٠	١٩٩٣	١٩٩٠	١٩٩٣
البرتقال	٤	٥	٣٠	٤٠
التفاح	٦	٨	٤٠	٥٠
الموز	٣	٤	٢٠	٣٠
العنب	٢	٣	١٠	٢٠
المانجو	٥	٦	١٠	٣٠

وسنستخدم الرموز التالية :
ع . سعر السلعة سنة الأساس .
ع ١ سعر السلعة سنة المقارنة .

ك . الكمية المستهلكة من السلعة سنة الأساس .

ك ١ الكمية المستهلكة من السلعة سنة المقارنة .

أولاً : الرقم القياسي التجميعي البسيط لأسعار الفاكهة .

$$= \frac{\text{مجـ ع}_1}{\text{مجـ ع}_0} \times ١٠٠$$

$$= \frac{٢٦}{٢٠} \times ١٠٠$$

= ١٣٠%

وذلك يعني أن أسعار الفاكهة إرتفعت من عام ٩٠ إلى ٩٣ بنسبة ٣٠% .

ثانياً : الرقم القياسي التجميعي البسيط للإنفاق على الفاكهة ويحسب هـذا الرقم كالتالي :

$$= \frac{\text{مجـ ع}_1 \text{ك}_1}{\text{مجـ ع}_0 \text{ك}_0} \times ١٠٠$$

$$= \frac{٩٦٠}{٤٩٠} \times ١٠٠$$

= ١٩٦%

وهذا يعني أن الإنفاق على الفاكهة قد إزداد بنسبة ٩٦% من سنة الأسـاس إلى سنة المقارنة .

ثالثاً : الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأســـاس . (لاسبير) .

$$= \frac{\text{مجـ ع}_{١}\text{ك}_{١}}{\text{مجـ ع . ك .}} \times ١٠٠$$

$$= \frac{٦٤٠}{٤٩٠} \times ١٠٠ = ١٣,٦\%$$

رابعاً : الرقم القياسي التجمعي للأسعار المرجح بكميات ســـنة المقارنـــة (باش) .

$$= \frac{\text{مجـ ع}_{١}\text{ك}_{١}}{\text{مجـ ع . ك}_{١}} \times ١٠٠$$

$$= \frac{٩٦٠}{٧٤٠} \times ١٠٠ = ١٢٩,٧\%$$

خامساً : الرقم القياسي الأمثل (فيشر)

$$= \sqrt{\text{رقم لا سبير} \times \text{رقم باشي}} \times ١٠٠$$

$$= \sqrt{\frac{\text{مجـ ع ١ ك .}}{\text{محـ ع . ب.}} \times \frac{\text{محـ ع ١ ك.}}{\text{محـ ع. ك ١}}} \times ١٠٠$$

$$= \sqrt{\frac{٦٤٠}{٤٩٠} \times \frac{٩٦٠}{٧٤٠}} \times ١٠٠$$

$$= \sqrt{١,٦٩٤٤} \times ١٠٠$$

$$= ١,٣٠ \times ١٠٠$$

$$= ١٣٠\%$$

الفصل الحادي عشر

مبادئ الاحتمالات

(١) مفهوم الاحتمالات :

كثيراً ما نتحدث عن احتمال ارتفاع درجة الحرارة في اليوم التالي واحتمال فوز فريق كرة قدم معين على فريق آخر . وأحياناً نجد أننا نعبر عن هذه الاحتمالات بتقدير عددي . كأن نقول أن احتمال سقوط أمطار غداً ٢٠% واحتمال وصول طائرة الخطوط البريطانية القادمة من لندن ٩٥% وهكذا .

وهذه التقديرات العددية للاحتمالات لا تستند إلى أساس رياضي لكن قد تعتمد على خبرات ومعلومات سابقة عن الطقس وعن تتبع لفترات طويلة لوصول طائرة الخطوط البريطانية القادمة من لدن .

وفي ضوء ذلك يمكن تعريف الاحتمالات على النحو التالي :
"نظرية الاحتمالات هي فرع من فروع الرياضة التطبيقية يهتم بدراسة تأثير الصدفة على الظواهر والأشياء" . لهذا لابد لنا من إيضاح كلمة "صدفة – chance" – هذه الكلمة التي تعودنا على سماعها في حياتنا اليومية ويمكن توضيح مفهومها على النحو التالي :
"إن لفظ صدفة وثيق الصلة بلفظ احتمال "Probability" وكلمة "احتمال" هي كلمة شائعة في لغتنا اليومية ودائماً نستعملها عندما نتكلم عن شئ يتحكم في حدوثه عوامل الصدفة" .

كذلك يمكن النظر إلى الاحتمالات على أنها أحد فروع الرياضيات الذي يهتم بدراسة نتائج التجارب أو المحاولات العشوائية . وتسمى التجربة أو المحاولة عشوائية إذا كانت نتائجها غير مؤكدة أي لا نستطيع التنبؤ بها .

فمثلاً ، إذا ألقيت قطعة معدنية من النقود فإننا لا نستطيع أن نتنبأ إذا كان السطح العلوي لها سيظهر صورة أو كتابة ، إذا فهذه محاولة أو تجربة عشوائية كذلك عند سحب ورقة عشوائياً من مجموعة أوراق اللعب المحكمة الخلط (الكوتشينة) فإننا لا نعلم إذا كانت الورقة المسحوبة ستظهر صورة أو عدداً ، إذا فهي محاولة عشوائية .

وعلى العموم فإن نتائج التجارب تنقسم إلى ثلاثة أنواع من وجهة نظر الاحتمالات هي كما يلي :

(أ) نتائج أو حوادث مؤكدة :

وهي نتائج لابد من وقوعها أو حدوثها .

مثال (١) : إذا ألقيت تفاحة في الهواء فإننا نعلم أنها لابد وأن تسقط على الأرض . هنا التجربة هي إلقاء التفاحة في الهواء ، والنتيجة هي سقوط التفاحة على الأرض .

مثال (٢) : إذا كان لدينا صندوق به ٨ كرات بيضاء اللون ، سحبت منه كرة واحدة فلابد أن تكون الكرة المسحوبة بيضاء .

هنا التجربة هي سحب كرة من الصندوق ، والنتيجة أن الكرة بيضاء : إذا فهذه نتيجة مؤكدة . وإذا كانت الحادثة مؤكدة الوقوع فإنه يقال أن احتمال وقوعها يساوي واحد .

أي أن احتمال سقوط التفاحة (في المثال ١) = ١

وكذلك احتمال أن تكون الكرة المسحوبة بيضاء في (المثال ٢) = ١

(ب) نتائج أو حوادث مستحيلة :

وهي تلك النتائج أو الحوادث المستحيل وقوعها .

مثال (٣) : هل يمكن سحب كرة حمراء من صندوق لا يحتوي إلا على كرات بيضاء .

التجربة هنا هي سحب كرة من الصندوق ، والنتيجة المطلوبة أن تكون الكرة حمراء ، إذاً فهذه حادثة مستحيلة .

مثال (٤) : أن يعيش شخص ما إلى الأبد ، هذه حادثة مستحيلة . وإذا كانت الحادثة مستحيلة الوقوع فإنه يقال أن احتمال وقوعها يساوي صفر .

أي أن احتمال سحب كرة حمراء من صندوق لا يحتوي إلا على كرات بيضاء (في المثال ٣) = صفر .

وكذلك احتمال أن يعيش شخص ما إلى الأبد (في المثال ٤) = صفر .

(ج) حوادث أو نتائج غير مؤكدة (محتملة أو ممكنة) :

وهي نتائج التجارب العشوائية التي ذكرناها سابقاً والتي لا نستطيع أن نتنبأ بوقوعها ، ولكننا نستطيع باستخدام تعريف الاحتمالات أن نحسب احتمال وقوعها .

حالات الاحتمالات :

(أ) الحالات المتماثلة (Eaqually Likely Cases)

هي تلك الحالات التي يكون لها فرص متكافئة من حيث الحدوث – أي لها نفس الفرصة .

فمثلاً لو كان لدينا صندوق به ١٠٠ كرة متشابهة في كل شئ عدا اللون منها ٥٠ كرة بيضاء ، و ٥٠ كرة سوداء – ورغبنا في سحب كرة من هذا الصندوق عشوائياً سنجد أن فرصة ظهور اللون الأبيض تعادل تماماً فرصة ظهور اللون الأسود ، وذلك بسبب تساوي أعداد الكرات من كل من اللونين ويعتبر اللونان في هذه الحالة حالتين متماثلتين . كذلك عند إلقاء قطعة عملة معدنية متزنة ومصنوعة من معدن متجانس وكانت عملية الإلقاء غير متحيزة فإن فرصة ظهور الصورة تعادل تماماً فرصة ظهور الكتابة وبهذا يمكن القول أن هاتين الحالتين (الصورة والكتابة) متماثلتين .

(ب) الحوادث الشاملة (Exhaustive Events)

يقال أن الحوادث أ١ ، أ٢ .... ، .... ، أن تشكل مجموعة من الحوادث الشاملة في تجربة معينة إذا كان لابد أن يتحقق واحد منها على الأقل عند إجراء التجربة ولا توجد نتيجة أخرى للتجربة تختلف عن هذه الحوادث . مثال ذلك عند إلقاء زهرة الطاولة فإن الأوجه الستة للزهرة (١-٢-٣-٤-٥-٦) تعتبر أحداثاً شاملة – كذلك عند إلقاء قطعة العملة يعتبر الوجهات (صورة ، كتابة) حدثين شاملين .

(ج) الحوادث المتنافية (Mutually Exclusive Events)

يقال أن الحوادث أ١ ، أ٢ ، ...... ، أن حوادث متنافية إذا استحال وجود أي أثنين (أو أكثر) منها في آن واحد .

فمثلاً في تجربة إلقاء زهرة الطاولة تعتبر الأوجه الستة حوادث متنافية لعدم إمكانية حدوث أي أثنين منها في آن واحد وكذلك في تجربة إلقاء قطعة العملة يعتبر الوجهان (صورة ، كتابة) حدثين متنافيين .

(د) الحالات الممكنة (Possible Cases)

هي مجموعة النتائج (أو الحالات) التي يمكن أن تنتج عند إجراء التجربة. فلو كانت التجربة هي إلقاء زهرة الطاولة مرة واحدة فإن الأوجه الستة للزهرة تعتبر هي الحالات الممكنة لهذه التجربة . كذلك إذا كانت التجربة هي سحب كرة واحدة من كيس يحتوي على عشرة كرات متماثلة فإن الحالات الممكنة تعتبر عشرة حالات متماثلة . وهكذا .

(ه) الحالات المواتية (Favorable Cases)

هي مجموعة النتائج التي تؤدي إلى تحقيق الحدث – وهي جزء من الحالات الممكنة للتجربة .

٣ – تعريف الاحتمالات :

يوجد للاحتمالات عدة تعاريف مختلفة نذكر منها تعريفين أثنين فقط واللذين لا يحتاجان إلى مفهوم رياضي متقدم هما :

أولاً : التعريف الكلاسيكي للاحتمالات .

ثانياً : التعريف التجريبي للاحتمالات .

أولاً : التعريف الكلاسيكي للاحتمالات :

إذا كنا بصدد إجراء تجربة ما فإن مجموعة النتائج التي يمكن أن تنتج عنها عددها ن من الحالات الشاملة المتنافية المتماثلة وكـــان م من هذه الحالات موات للحدث أ فإن احتمال وقوع الحدث أ يعرف بأنه النسبة $\frac{م}{ن}$.

فإذا رمزنا لاحتمال وقوع الحدث أ بالرمز ح (أ) فـيمكن كتابـــة هــذا الاحتمال في الصورة التالية :

ح (أ) = عدد الحالات المواتية للحدث أ \ عدد الحالات الممكنة للتجربة .

فمثلاً عند سحب ورقة من مجموعـــة أوراق اللعـــب المحكمـــة الخلـــط (الكوتشينة) نجد أن لدينا ٥٢ حالة متنافية ومتماثلة هي الحالات الممكنة للتجربة . فإذا كان الحدث أ هو الحصول على صورة يكون أمامنـــا ١٢ حالة مواتية لوقوع الحدث أ وهي عدد الصور في الكوتشينة وعلى هــذا يكون احتمال وقوع الحدث أو مساوياً $\frac{١٢}{٥٢}$ ، وتكتب في صورة رمزية كما يلي :

$$ح (أ) = \frac{١٢}{٥٢} = \frac{٣}{١٣}$$

كذلك إذا كانت التجربة هي إلقاء زهرة نرد متزنة تكون الحالات الممكنة لهذه التجربة ٦ حالات شاملة ومتنافية ومتماثلة . فإذا كان الحــدث هــو الحصول على عدد زوجي من النقاط فإن الحالات المواتية لهذا الحــدث هي ٣ حالات (٢-٤-٦) وهي الأوجه التي تحمل عدداً زوجياً من النقط وبهذا يكون احتمال وقوع هذا الحدث مساوياً $\frac{٣}{٦}$ - وتكتب : ح (أ) = $\frac{٣}{٦}$ = $\frac{١}{٢}$ وهكذا .

مثال (٥) : عند إلقاء زهرة نرد مرة واحدة – ما هو احتمال أن يكون مجموع النقط على السطح العلوي لها أكثر من ٤ ؟

التجربة : هي إلقاء زهرة النرد .

إذا الحالات الممكنة هي : ن = ٦ حالات متماثلة .

الحدث أ هو : أن يكون مجموع النقط على السطح العلوي للزهرة أكبر من ٤ .

إذا الحالات المواتية هي : م = ٢ (وهي الحالتين ٥ ، ٦) .

إذاً ح (أ) $\frac{م}{ن} = \frac{٢}{٦} = \frac{١}{٣}$.

مثال (٦) . عند إلقاء قطعتي عملة مرة واحدة (أو إلقاء قطعة واحدة مرتين متتاليتين) ما هو احتمال الحصول على صورتين ؟

الحـــــل

التجربة هي : إلقاء قطعتي عملة .

الحالات الممكنة هي : ن = ٢×٢ = ٤ حـــالات .

وذلك لأن القطعة الأولى لها وجهان كل وجه منهما يمكن أن يناظره وجهان للقطعة الثانية . وهذه الحالات الأربع يمكن حصرها لو رمزنا للصورة بالرمز ص والكتابة بالرمز ك كما يلي :

ص ص – ص ك – ك ص – ك ك .

الحدث أهو : الحصول على صورتين .

إذا الحالات المواتية هي : حالة واحدة وهي (ص ص) .

إذا ح (أ) = $\frac{\text{عدد الحالات المواتية}}{\text{عدد الحالات الممكنة}}$ $\frac{١}{٢}$

مثال (٧) : عند إلقاء زهرتين متزنتين من زهرات النرد مرة واحدة (أو إلقاء زهرة واحدة مرتين) ما هو احتمال أن يكون مجموع النقط على السطحين العلويين :

: مساوياً ٩ ؟ : ٩ فأكثر ؟

الحـــل

التجربة : إلقاء زهرتي نرد متزنتين :

ن = ٣٦ حالة متماثلة (٦ حالات للزهرة الأولى كل حالة منها يقابلها ٦ حالات للزهـــرة الثانية وبذلك يكون عدد الحالات الممكنة مساوياً ٦×٦= ٣٦ حالة) .

ويمكن حصر الحالات الممكنة في الشكل التالي :

نرمز للزهرة الأولى بالرمز س والثانية بالرمز ص

ص / س	١	٢	٣	٤	٥	٦
١	(١,١)	(٢,١)	(٣,١)	(٤,١)	(٥,١)	(٦,١)
٢	(١,٢)	(٢,٢)	(٣,٢)	(٤,٢)	(٥,٢)	(٦,٢)
٣	(١,٣)	(٢,٣)	(٣,٣)	(٤,٣)	(٥,٣)	(٦,٣)
٤	(١,٤)	(٢,٤)	(٣,٤)	(٤,٤)	(٥,٤)	(٦,٤)
٥	(١,٥)	(٢,٥)	(٣,٥)	(٤,٥)	(٥,٥)	(٦,٥)
٦	(١,٦)	(٢,٦)	(٣,٦)	(٤,٦)	(٥,٦)	(٦,٦)

الحالات السابقة تمثل ٣٦ نتيجة – فمثلاً النتيجة (٥ ، ٢) معناهـا أن الزهـرة الأولى نتيجتها الوجه الذي عليه ٥ فقط والزهرة الثانية الوجـه الـذي عليـه نقطتين.

إذا الحدث $أ_١$: هو أن يكون مجمع النقط على السطحين العلويين ٩ .

إذا الحالات المواتية $م_١$: هي تلك الحالات التي تبدو بين الخطين المائلين فـي الجدول السابق وعددها ٤ حالات .

أي أن $م_١ = ٤$

$$إذا\ ح(أ_١) = \frac{م_١}{ن} = \frac{٤}{٣٦} = \frac{١}{٩}$$

إذا الحدث $أ_٢$: هو أن يكون مجموع النقط على السطحين العلويين ٩ فأكثر .

الحالات المواتية م ٢ : هي تلك الحالات الموجودة بين الخطين المـائلين فـي الجدول السابق بالإضافة إلى كل الحالات الموجودة أسفل هـذين الخطـين لأن كلها تحقق الحدث المطلوب $أ_١$ أي أن مجموع كل منها إما ٩ أو أكثـر مـن ٩ وعددها ١٠ حالات متماثلة – أي أن :

$م_٢ = ١٠$

$$إذا\ ح(أ_٢) = \frac{م_٢}{ن} = \frac{١٠}{٣٦} = \frac{٥}{١٨}$$

٤ – المبادئ الأولية للاحتمالات

مما سبق نستنتج ما يلي :

(I) صفر $\leq$ ح (أ) $\leq$ ١ .

(II) ح (أ) = صفر إذا كانت أ حادثة مستحيلة .

(III) ح (أ) = ١ إذا كانت أ حادثة مؤكدة

(IV) إذا كان احتمال وقوع الحادثة أ هو ح وكان احتمال عدم وقوعها هو ل فإن ح + ل = ١ .

أي أن احتمال وقوع أي حادثة + احتمال عدم وقوعها = ١ .

وذلك لأنه من المؤكد أن تقع الحادثة أو لا تقع .

٤ - ١ بعض قوانين الاختيار الهامة :

لإمكان حل مسائل الاحتمالات فإننا سنعتمد على بعض قوانين موضوع الاختيار ونذكر منها على الأخص :

(أ) عدد الطرق التي يمكن بها اختيار س من الأشياء من بين ن من هذه الأشياء.

$$= {}^{ن}ق_{س} = \frac{ن\,!}{س\,!\,(ن - س)\,!}$$

حيث أن :

ن ! = ن (ن - ١) (ن - ٢) × .... × ٢ × ١

فمثلاً : ٤ ! = ٤ × ٣ × ٢ × ١ = ٢٤

مثال (٨) : إذا كان لدينا ٤ رجال وأريد إرسال بعثة منهم مكون من رجلين ، فإنه يمكن اختيار أعضاء هذه البعثة بعدد من الطرق مساوياً ${}^{٤}ق_{٢}$ طريقة .

$$\text{حيث } {}^{٤}ق_{٢} = \frac{٤\,!}{٢\,!\,(٤ - ٢)\,!} = \frac{٤\,!}{٢\,! \times ٢\,!}$$

$$= \frac{٤ \times ٣ \times ٢ \times ١}{٢ \times ١ \times ٢ \times ١} = ٦ \text{ طرق .}$$

مثال (٩) : صندوق به ٨ كرات متماثلة ، سحبت منه ٣ كرات ، **فما هي** عدد الطرق التي يمكن بها إجراء هذه العملية ؟ .

الحـــل

عدد الطرق = $^{٨}ق_{٣}$ = $\frac{٨!}{٣! \times ٥!}$ = $\frac{٨ \times ٧ \times ٦ \times ٥!}{٣ \times ٢ \times ١ \times ٥!}$ = ٥٦ طريقة

(ب) إذا أمكن إجراء عملية ما بطرق مختلفة عددها ل وأمكن إجــراء عمليــة أخرى بطرق مختلفة عددها م ، فإن عدد الطرق التي يمكن بها إجراء العمليتين معا هو ل م .

مثال (١٠) : ما هو عدد ا لطرق التي يمكن بها تكوين بعثة من ٣ رجال ، ٢ نساء من بين ٦ رجال ، ٥ نساء .

الحـــل

عدد طرق اختيار الرجال = $^{٦}ق_{٣}$ = $\frac{٦!}{٣! \times ٣!}$ = $\frac{٦ \times ٥ \times ٤}{٣ \times ٢ \times ١}$ = ٢٠ طريقة

عدد طرق اختيار النساء = $^{٥}ق_{٢}$ = $\frac{٥!}{٢! \times ٣!}$ = $\frac{٥ \times ٤}{٢ \times ١}$ = ١٠ طرق

عدد طرق تكوين البعثة = ١٠ × ٢٠ = ٢٠٠ طريقة .

٥ – أمثلة على الاحتمالات :

مثال (١١) : صندوق به ٥ كرات بيضاء ، ٤ حمراء ، سحبت منه كرة واحدة فما هو احتمال أن تكون الكرة المسحوبة :

(I) بيضـــاء ؟ (II) حمـــراء ؟

الحـــــل

(I) عدد الطرق التي يمكن بها سحب كرة من الصندوق .

$$= {}^{9}ق_{1} = \frac{9!}{1! \times 8!} = 9 \text{ طرق}$$

– عدد الطرق التي يمكن بها سحب كرة بيضاء من الصندوق

$$= {}^{5}ق_{1} = \frac{5!}{1! \times 4!} = 5 \text{ طرق}$$

– عدد الطرق التي يمكن بها سحب كرة حمراء من الصندوق

$$= {}^{4}ق_{1} = \frac{4!}{1! \times 3!} = 4 \text{ طرق}$$

إذا (I) احتمال سحب كرة بيضاء = $\frac{5}{9}$

(II) احتمال سحب كرة حمراء = $\frac{4}{9}$

يلاحظ في هذا المثال أن مجموع الاحتمالات فـــي (I) ، (II) يســـاوي الواحد الصحيح لأن الكرة المسحوبة إما أن تكون بيضاء أو حمراء وهذه حادثة مؤكدة .

مثال (١٢) : صندوق به ٥ كرات بيضاء ، ٤ حمراء ، سحبت منه كرتـان ، فما احتمال أن تكون الكرتان :

(I) بيضـاء ؟ (II) واحدة بيضاء والأخرى حمراء ؟

(I) عدد الطرق التي يمكن بها سحب كرتان من الصندوق :

$= {}^{9}ق_{2} = \frac{9!}{2! \times 7!} = 36$ طريقة

عدد الطرق التي يمكن بها سحب كرتين بيضاء من الصندوق :

$= {}^{5}ق_{2} = \frac{5!}{2! \times 3!} = 10$ طريقة

إذا احتمال أن تكون الكرتان بيضاء $= \frac{10}{36}$

(II) عدد الطرق التي يمكن بها سحب كرة بيضاء من الصندوق :

$= {}^{5}ق_{1} = 5$ طرق

إذا عدد الطرق التي يمكن بها سحب كرة حمراء من الصندوق

$= {}^{4}ق_{1} = 4$ طرق

إذاً عدد الطرق التي يمكن بها سحب كرة بيضاء وكرة حمـراء مـن الصندوق $= 5 \times 4 = 20$ طريقـة .

إذا احتمال أن تكون الكرتان واحد بيضاء والأخرى حمراء $= \frac{20}{36}$

مثال (١٣) : ألقيت زهرة نرد مرة واحدة ، فما هو احتمال أن يكون السـطح العلوي لها : (I) أقل من ٣ ؟ (II) ٤ فأكثر .

الحــــل

(I) عدد الطرق التي يمكن أن يظهر بها السطح العلوي = ٦
وعدد الطرق التي يمكن أن يظهر بها على السطح رقم أقل من ٣
= ٢ (وهما ظهور الوجه ١ أو الوجه ٢) .
إذاً إحتمال (الحصول على عدد أقل من ٣) = $\frac{٣}{٦} = \frac{١}{٢}$

(II) عدد الطرق التي يمكن أن يظهر بها على السطح العلوي ٤ فـأكثر
= ٣ (وهي الوجه ٤ أو الوجه ٥ أو الوجه ٦) .
إذاً احتمال (الحصول على عدد ٤ فأكثر) = $\frac{٢}{٦} = \frac{١}{٢}$

مثال (١٤) : ألقيت زهرتي نرد مرة واحدة ، فما هو احتمال أن يكون مجموع النقط على السطح العلوي لهما ٥.

الحــــل

عدد الطرق التي يمكن أن يظهر بها السطح العلوي للزهرة الأولى = ٦ طرق .
عدد الطرق التي يمكن أن يظهر بها السطح العلوي للزهرة الثانية = ٦ طرق .
عدد الطرق التي يمكن أن يظهر بها السطح العلوي للزهرتين معـاً = ٦×٦ = ٣٦ طريقة .
عدد الطرق التي يمكن أن يكون مجموع النقط على السطح العلوي ٥ هو = ٤ طرق .
وهى (١ ، ٤) أو (٢ ، ٣) أو (٣ ، ٢) أو (٤ ، ١)
إذا احتمال (أن يكون المجموع ٥) = $\frac{٤}{٣٦} = \frac{١}{٩}$

٦ – التعريف التجريبي للاحتمال :

إذا كررنا تجربة معينة مرات عددها ن (تحت نفس الظروف) ولاحظنا أن حادثاً معيناً أ قد تحقق في م من هذه المرات فإن النسبة $\frac{م}{ن}$ تسمى بـــالتكرار النسبي للحدث أو تعتبر قيمة تجريبية لاحتمال وقوع هذا الحدث وتقترب هـــذه القيمة التقريبية من احتمال وقوع الحدث أ كلما كبرت ن حتى أنه عندما تصبح ن كبيرة كبراً لا نهائياً تصبح هذه القيمة التقريبية هي احتمال وقوع الحـــدث أ ويكون :

احتمال وقوع الحدث أ هو ح (أ) = $\frac{م}{ن}$ عندما تكبر ن كبراً لا نهائياً .

ح (أ) = نهـــا $\frac{م}{ن}$

ن ⟵ ما لا نهاية

(حيث أن نها هي اختصار لكلمة نهاية – ويكون معنى الرمز السابق أنه فـــي النهاية عندما تكبران كبراً لا نهائياً يكون الاحتمال ح (أ) = $\frac{م}{ن}$) .

هذا التعريف يقوم على أساس ملاحظة الظاهرة موضوع الدراسة فمن المعلوم أن أي ظاهرة طبيعية مثل المواليد والوفيات وغير ذلك من الظـــواهر تخضـــع لصفة نظامية محددة هذه الصفة النظامية لا تظهر في الحالات القليلـــة العـــدد ولكنها تظهر بوضوح في الحالات الكبيرة العدد . كما أن هذا التعريف يســـمى التعريف التجريبي لأنه يعتمد على ملاحظة التجربة – كما أنه يســـمى أحيانـــاً بالتعريف البعدي لأن الاحتمال يتم حسابه بعد إجراء التجربة وهذا يختلف عن التعريف الكلاسيكي الذي يمكن استخدامه في حساب الاحتمـــال قبـــل إجـــراء التجربة .

مثال (١٥) : لدينا قطعة عملة معروف أنها غير متزنة – تم إلقاؤها ألف مـــرة فظهرت الصورة ٥٥٠ مرة والكتابة ٤٥٠ مرة ، فما هو احتمال الحصول على الصورة عند إلقاء هذه القطعة ؟

الحـــل

بما أن الصورة ظهرت ٥٥٠ مرة من بين ألف مرة – إذن نسبة ظهور الصورة = $\frac{٥٥٠}{١٠٠٠}$ = ٥٥ر وهذه النسبة يمكن اعتبارها تقريباً احتمال الحصـــول علـــى الصورة وبالتالي يكون الاحتمال المطلوب هـــو : ح = ٠,٥٥

مثال (١٦) : أجرى طبيب ٥٠٠ عملية جراحية ونجح منها ٤٨٠ عملية فمـــا احتمال نجاح عملية يجريها هذا الطبيب ؟

الحـــل

عدد مرات إجراء العملية ن = ٥٠٠

عدد مرات نجاح العملية م = ٤٨٠

احتمال نجاح العملية = $\frac{٤٨٠}{٥٠٠}$ = ٠,٩٦

مثال (١٧) : في مصنع للمصابيح الكهربية تبين أن من بين كل ١٠٠٠ مصباح يوجد ٥٠ مصباحاً غير صالح للاستعمال . فما هو احتمال وجود مصباح جيد؟

الحـــل

عدد المصابيح ن = ١٠٠٠ مصباح

عدد المصابيح الجيدة م = ٩٥٠ مصباحاً

إذا الاحتمال المطلوب = $\frac{م}{ن}$ = $\frac{٩٥٠}{١٠٠٠}$ = ٠,٩٥

الفصـــل الثاني عشـــر

التوقع الرياضــــــي

Mathematical Expectation

تتطلب الحياة التجارية في كثير من الحالات استخدام الاحتمالات فـــي ترشيد القرارات لاختيار أفضلها .

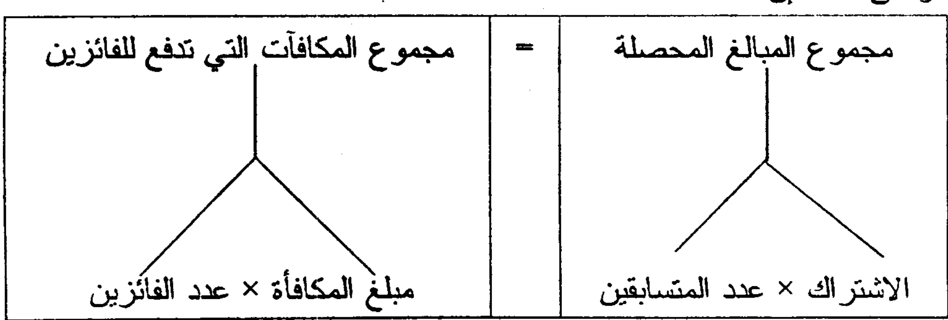
ويعتبر التوقع الرياضي للكسب أو الخسارة هو المؤثر الأكثر دقة عند إجراء هذا التفضيل . فإذا فرضنا أن هناك منظم لإحدى المسابقات حيث يقـــوم بجميع الاشتراكات من مجموع المتسابقين على أن تكفـــي هـــذه الاشـــتراكات بالكامل لسداد مكافآت الفوز للأشخاص الفائزين ، فبفرض أن :

عدد المتسابقين = ن

عدد الأشخاص الفائزين = م

قيمة المكافأة أو الرهان = س

وعلى ذلك فإن :ـــ



$$ت(س) \times ن \quad = \quad س \times م$$

$$ت(س) \quad = \quad س \times \frac{م}{ن}$$

ت س = س × ح (س)

ويطلق على ت (س) التوقع الرياضي أو القيمة المتوقعة أو المتوسطة للفوز وعلى ذلك فإن :

التوقع الرياضي = مبلغ المكافأة أو الرهان × احتمال الفوز بهذا المبلغ

التوقع الرياضي إذا كان مبلغ الرهان يدفع في تاريخ مستقبلي أو لاحق لتاريخ تحصيل الاشتراكات .

إذا تعهد المنظم بسداد مكافآت الفوز مستقبلاً بعد فترة من تحصيل الاشتراكات ويفرض أن المبالغ المحصلة (مبالغ الاشتراكات) سوف تستثمر لحساب مجموع المتسابقين بمعدل فائدة مركبة ع وحتى تاريخ سداد المكافأة وهذا يعني أن :

جملة الاشتراكات المحصلة في تاريخ سداد المكافآت = مجموع المكافآت المدفوعة للفائزين

ت (س) $\{$١ + ع$\}^{ن}$ = س × ح (س)

ت (س) = س × ح (س) × $\frac{١}{(١ + ع)^{ن}}$

ت (س) = س × ح (س) × $ح^{ن}$

وتمثل ح ن القيمة الحالية للجنية ع % ولمدة ن من السنوات وتستخدم الجداول المالية للفائدة المركبة في حساب قيم ح ن وتستخدم الصيغة السابقة في حساب أقساط التأمين التي تحصل من مجموعة المؤمنين مقدماً عند التعاقد في مقابل سداد مبلغ التأمين عند تحقق الخطر المؤمن منه .

القيمة المتوقعة Expected Value

تعرف القيمة المتوقعة بأنها مجموع التوقع الرياضي لجميع القيم التي يمكن أن يأخذها المتغير العشوائي فإذا رمزنا للقيمة المتوقعة بالرمز ت (س)

فإن ت (س) = $س_١$ × ح ($س_١$) + $س_٢$ × ح ($س_٢$) + $س_٣$ × ح ($س_٣$)

+ .... + س ن × ح (س ن)

$$ت (س) = \sum_{ر = ١}^{ر = ن} س_ر \times ح(س_ر)$$

وتستخدم هذه القيمة في اتخاذ القرارات الإدارية بناءاً على النتائج المنبثقة عن التجربة فإذا كانت القيمة الناتجة موجبة يمكن القول أن هذا القرار يجب الإقدام على اتخاذه ، أما إذا كانت سالبة يمكن القول أن هذا القرار يجب عدم اتخاذه ، أما إذا كانت القيمة صفر يمكن القول أن المسابقة عادلة .

وفيما يلي بعض الأمثلة التي توضح ذلك :

مثال ١ :

تاجر ربح ١٠٠٠٠ جنيه في الشهر خلال سنة ١٩٩٨ حيث كان الطلب ممتاز وربح ٦٠٠٠ جنيه حيث كان الطلب معتدلاً ، وخسر ١٠٠٠ جنيه حيث كان الطلب ردئ ، فإذا كان احتمالات تحقق صور الطلب السابقة على الترتيب كما يلي ٠.٧ ، ٠.٥ ، ٠.٣. فما هي القيمة المتوقعة لربح هذا التاجر في الشهر في سنة ٢٠٠٠ .

الحـــل

ت (س) = س$_{١}$ + ح (س$_{١}$) + س$_{٢}$ × ح (س$_{٢}$) + س$_{٣}$ × ح (س$_{٣}$)

= ١٠٠٠٠ × .٧ +٦٠٠٠×.٥ + (١٠٠٠٠×.٣)

= ٧٠٠٠ + ٣٠٠٠ − ٣٠٠

= ١٠٠٠٠ − ٣٠٠ = ٩٧٠٠ جنيه

مثال ٢ :

تعهد شخص بسداد مبلغ ١٠٠ جنيه لكل متسابق يرمي مكعب منـــتظم ويكون على السطح العلوي للمكعب أحد الأقسام الفردية على أن يدفع المتسابق مبلغ ٨٠ جنيه إذا ظهر على السطح العلوي للزهرة الرقم (٥) . إحسب القيمـــة المتوقعة لفوز شخص بطلب الاشتراك في هذه المسابقة .

الحـــل

القيمة المتوقعة للفوز = ١٠٠ × $\frac{٣}{٦}$ − ٨٠ × $\frac{١}{٦}$

= ٥٠ − ١٣,٣٣

= ٣٦,٦٧

مثال ٣ :

إحسب القيمة المتوقعة للمتغير العشوائي التالي :ـــ

س	١	٢	٣	٤	٥	٦
ح (س)	$\frac{١}{٦}$	٤/[illegible]	٣/[illegible]	$\frac{٢}{٦}$	$\frac{٥}{٦}$	$\frac{١}{٦}$

الحــــل

ت (س) = $س_١$ + ح ($س_١$) + $س_٢$ × ح ($س_٢$) + $س_٣$ × ح ($س_٣$)

$س_٤$ + ح ($س_٤$) + $س_٥$ ×ح ($س_٥$) + $س_٦$ × ح ($س_٦$)

$$= ١ \times \frac{١}{١٦} + ٢ \times \frac{٤}{١٦} + ٣ \times \frac{٣}{١٦} + ٤ \times \frac{٢}{١٦} + ٥ \times \frac{٥}{١٦} + ٦ \times \frac{١}{١٦}$$

$$= \frac{١}{١٦} + \frac{٨}{١٦} + \frac{٩}{١٦} + \frac{٨}{١٦} + \frac{٢٥}{١٦} + \frac{٦}{١٦}$$

$$= \frac{٥٧}{١٦}$$

مثال ٤ :

تعاقد شخص عمره ٣٠ سنة مع إحدى شركات التأمين على أن تلتــزم الشركة بدفع مبلغ ١٠٠٠ جنيه للشخص إذا ظل على قيد الحياة حتى تمام السن ٤٠ فإذا علمت أن من بين كل ١٠٠٠٠ شخص في تمام السن ٣٠ يعيش منهم ٨٠٠٠ شخص عند تمام السن ٤٠ إحسب القسط الوحيد الصافي لهذا العقــد إذا علم أن ح ١٠ ٤% = ٦٧٥٥٦.

الحــــل

القسط الوحيد الصافي = مبلغ التأمين × احتمال الحياة × $^{ح}ن_{ع\%}$

$$= ١٠٠٠٠ \times \frac{٨٠٠٠}{١٠٠٠٠} \times .٦٧٥٥٧$$

= ٨٠٠٠ × ٦٧٥٥٦.

= ٥٤٠٤,٤٨ جنيه

القيمة المتوقعة في توزيع ثنائي الحدين :

القيمة المتوقعة في هذا التوزيع = عدد المحاولات × احتمال النجاح .

أي أن ت (س) = ن × ح

ويمكن حساب القيمة المتوقعة في توزيع ثنائي الحدين بإحدى الطريقتين الآتيتين :

الطريقة الأولى : ت (س) = $\sum_{ر = ١}^{ر = ن}$ س ر ح (س ر)

الطريقة الثانية : ت (س) = ن × ح

مثال (٥) :

ألقى شخص قطعة نقود على سطح أملس ثلاث مرات متتالية كون جدول توزيع ثنائي الحدين على أساس عدد مرات ظهور الصورة في كل رمية ثم إحسب القيمة المتوقعة لعدد مرات ظهور الصورة في الرميات الأربعة بطريقتين مختلفتين :

الحل

احتمال ظهور الصورة = $\frac{١}{٢}$ احتمال عدم الظهور = $\frac{١}{٢}$

ويتحدد التوزيع الاحتمالي لعدد مرات الصورة من خلال جدول توزيع ثنائي الحدين كالتالي :ـ

جدول توزيع نتائج الحدين

عدد مرات ظهور الصورة (ر)	قيمة الاحتمال ح (ر)
٠	$ح(٠) = {}^{٣}ق_{٠} \left(\frac{١}{٢}\right)^{٠} \left(\frac{١}{٢}\right)^{٣} = \frac{١}{٨}$
١	$ح(١) = {}^{٣}ق_{١} \left(\frac{١}{٢}\right)^{١} \left(\frac{١}{٢}\right)^{٢} = \frac{٣}{٨}$
٢	$ح(٢) = {}^{٣}ق_{٢} \left(\frac{١}{٢}\right)^{٢} \left(\frac{١}{٢}\right)^{١} = \frac{٣}{٨}$
٣	$ح(٣) = {}^{٣}ق_{٣} \left(\frac{١}{٢}\right)^{٣} \left(\frac{١}{٢}\right)^{٠} = \frac{١}{٨}$
٣	**مجموع الاحتمالات = ١**

وتتحدد القيمة المتوقعة لعدد مرات ظهور الصورة في الرميات الأربع بالطريقتين الآتيتين :

الطريقة الأولى : $ت(س) = \sum_{ر=١}^{ر=ن} س_{ر}\, ح(س_{ر})$

$$= ٠ \times \frac{١}{٨} + ١ \times \frac{٣}{٨} + ٢ \times \frac{٣}{٨} + ٣ \times \frac{١}{٨}$$

$$= ٠ + \frac{٣}{٨} + \frac{٦}{٨} + \frac{٣}{٨}$$

$$= \frac{١٢}{٨} = ١,٥ \text{ مرة}$$

الطريقة الثانية :- ت (س) = ن × ح

= ٣ × $\frac{١}{٢}$ = ١,٥ مرة .

الفصل الثالث عشر

التباديل والتوافيق

Permutation and Combination

تفيد التباديل والتوافيق في معرفة بعض التوزيعات الاحتمالية وأهمهـا التوزيع الثنائي (ذات الحدين) . كما تعتبر أحد الأساليب الرياضية التي يمكـن استخدامها في تقدير عدد مرات حدوث حدث معين .

التباديل Permutation

يمكن تعريف التباديل على أنها ترتيب معين لمجموعة من الأشـياء ، وعندما يتغير هذا الترتيب نحصل على تبديل آخر .

مثال :

ما هو عدد الطرق التي يمكن بها ترتيب أربعة أفراد A,B,C,D ؟

الحـــل :

A, B, C, D	B, A, C, D	C, D, A, B	D, A, B, C
A, B, C, D	B, A, D, C	C, D, B, A	D, A, C, B
A, B, D, C	B, C, D, A	C, B, A, D	D, B, C, A
A, C, D, B	B, C, A, D	C, B, D, A	D, B, A, C
A, D, B, C	B, D, A, C	C, A, B, D	D, C, B, A
A, D, C, B	B, D, C, A	C, A, D, B	D, C, A, B

أي أن عدد الطرق (التباديل) التي يمكن بها ترتيب أربعة أفراد = ٢٤ ويمكـن الحصول على عدد التباديل من واقع المعادلة التالية :

$$ {}_{n}p_{r} = \frac{n!}{(n-r)!} $$

$$4p4 = \frac{4!}{0!} = \frac{(4)(3)(2)}{1} = 24$$

$$_np_n = n! \qquad \text{When : } n - r$$

التوافيق Combination

تعتبر التوافيق تجميع للأشياء بغض النظر عن الترتيب ، وهي في الواقع العملية الأولى التي تقوم بها ، أي هي عملية الاختيار أو الانتخاب .

مثال :

إذا كان هناك أربعة أفراد ونريد اختيار أثنين منهم لإرسالهما في بعثة . والمطلوب هو إيجاد عدد الطرق التي يمكن بواسطتها اختيار أفراد العينة .

الحـــل :

بفرض أن الأشخاص الأربعة هم : A, B, C, D

فإنه يمكن تشكيل البعثة على الوجه التالي :

A, B   A, C   A, D

B, C   B, D   C, D

عدد الطرق (التوافيق) = ٦ طرق

ويمكن الوصول إلى هذه النتيجة من خلال المعادلة التالية للتوافيق :

$$_nC_r = \frac{n!}{r i (n-1)}$$

$$_4C_2 = \frac{4!}{2!\ 2!} = \frac{(4)\ (3)\ (2)\ (1)}{(2)\ (1)\ (2)\ (1)} = \boxed{٦}$$

وللمقارنة يمكن تقدير التباديل التي يمكن بها ترتيب أثنين من الأربعـــة أفــراد المرشحين للبعثة :

$$_{n}p_{r} = \frac{n\,!}{(n-r)\,!}$$

$$_{4}p_{2} = \frac{4\,!}{2\,!} = \frac{(4)\ (3)\ (2)\ (1)}{(2)\ (1)} = \boxed{12}$$

ومن الواضح أن عدد التوافيق < عدد التباديل وذلك لأننا في التوافيـــق عنـــدما نقول نرسل (A , B) في بعثة ، فإننا لا نغير من الأمر شيئاً إذ يمكنا أن نرسل (B, A) في البعثة .

على عكس الحال في التباديل فلو أردنا مثلاً منح جائزتين ، جائزة أولى وجائزة ثانية لكان هناك فارقاً بين (B,A) و (A,B) ويرجع ذلك إلى أنه عندما نقـــوم بعملية التباديل لعدد من الأشياء فإننا في الواقع نقوم بعمليتين منفصلتين علـــى هذا العدد من الأشياء :

١- عملية اختيار الأشياء .

٢- عملية ترتيب هذه الأشياء .

مثال :

إذا كان لك خمسة أصدقاء فبكم طريقة يمكن دعوة ثلاثة منهم لتنـــاول الغذاء بفرض وجود أثنين منهم لا يمكن الجمع بينهما .

بفرض أن الأصدقاء الخمسة هم : A, B, C, D, E وأنـــه لا يمكن الجمع بين D, E ، ففي هذه الحالة يمكن :

١- استبعاد (D, E) .

٢- استبعاد D ودعوة E .

٣- استبعاد E ودعوة D .

يطبق قانون التوافيق :

$$_{n}C_{r} = \frac{n\,!}{r\,!(n-r)\,!}$$

(١) استبعاد (D, E) $_{3}C_{3} = 1$

(٢) استبعاد D ودعوة E $_{3}C_{2} = 3$

(٣) استبعاد E ودعوة D $_{3}C_{2} = 3$

إجمالي عدد الطرق = ١ + ٣ + ٣ = ٧ طـــــرق .

الفصل الرابع عشـــر

تحليل التباين

Analysis of Variation

تعتبر طريقة تحليل التباين من الطرق الإحصائية المستخدمة في تقدير درجة دقة النتائج أو بمعنى آخر درجة معنوية النتائج ، أي أنها تمكننـــا مـــن الحكم على درجة مطابقة مجموعة الإحصاءات المتحصل عليها من البحـــث أو التجربة لما تتوقعه بناء على النظرية الفرضية الموضوعة قبل إجراء البحث .

النظرية الفرضية Null Hypothesis

يطلق على النظرية الفرضية الإحصائية فرض العدم كما أسماها العالم Fisher ، وهي تفترض أن الفروق بين متوسطات المعاملات المختلفة للتجربة هي فروق عشوائية Random Sampling Error أو ما يمكن تسميتها بالخطأ التجريبي Experimental Error وهذه الفروق العشوائية ناتجـــة عـــن أخـــذ عينات مختلفة من مجتمع واحد وأنها ليست ناتجة عن تأثير المعاملات .

وتهدف اختبارات المعنوية بصفة عامة إلى محاولة التعرف على ما إذا كانت الفروق بين العينات أو المعاملات هي فروق عشوائية بين عينات المجتمع الواحد كما تفترض النظرية الفرضية ، أم أن هذه الفروق هي في الحقيقة فروق أساسية ناتجة عن تأثير المعاملات المختلفة ، وأن تلك العينات مـــأخوذة مـــن مجتمعين مختلفين ، وفي هذا الحالة نرفض النظرية الفرضية ويطلـــق علـــى الفروق الناتجة عن تأثير المعاملات عادة أسم الفروق المعنوية Significant Differences .

الأساس الذي يقوم عليه اختبار المعنوية :

يتمثل الأساس الذي يعتمد عليه اختبار المعنوية في مقارنة الفرق بين العينات أو المعاملات في التجربة بالفروق بين العينات العشوائية المأخوذة من مجتمع واحد . وكما سبق الذكر فإن الفروق بين متوسطات العينات العشوائية ومتوسطات المجتمع تحدث كنتيجة لأخذ هذه العينات عشوائياً من مجتمع واحد، وأن قيمة هذه الفروق تعتمد على مدى الاختلافات الموجودة بين أفراد المجتمع بحيث تزداد قيمتها بزيادة الاختلافات وتقل كلما زاد تجانس المجتمع ، وأن الفروق بين متوسطات عينات نفس المجتمع تتوزع توزيعاً طبيعياً حول متوسط هذه الفروق والذي يساوي صفراً .

ولما كانت الفروق التي تحدث بين متوسطات العينات العشوائية المأخوذة من مجتمع واحد يمكن تقديرها بحساب الانحراف القياسي لتلك المتوسطات فإن مقارنة الفروق بين متوسطات المعاملات في التجربة بهذا الانحراف القياسي للمتوسطات يعطى فكرة عن مدى دلالة هذا الفرق .

فإذا كان الفرق بين هذه المتوسطات يساوي أو يقل عن الانحراف القياسي للمتوسطات دل ذلك على أن هذا الفرق يدخل ضمن الفروق العشوائية ، ومن ثم تقبل النظرية الفرضية . أما إذا كانت الفروق بين متوسطات المعاملات أكبر من الفروق العشوائية ، فإن النظرية ترفض لأن الفروق هنا ناتجة عن تأثير المعاملات .

استخدام تحليل التباين :

يستخدم تحليل التباين في النواحي التالية :

١- تقدير درجة الاعتماد على النتائج أو دقتها عند تقسيم ظاهرة ما إلى مجموعات نتيجة لتغير عنصر واحد في المتوسط .

٢- تقدير درجة معنوية النتائج في حالة دراسة أثر أكثر من متغير واحد على التغير في الظاهرة موضوع التحليل .

مستوى المعنوية Significance Level

هو درجة الاحتمال الذي تقبل أو ترفض على أساسه النظرية الفرضية والمتبع في الدراسات الاقتصادية هو استعمال مستويين للمعنوية هما :

(أ) مستوى المعنوية ٠١ر (١%) : يعني أن احتمال وقوع مشاهدة ما في المدى (U + 36) هو ٩٩% واحتمال وقوع المشاهدة خارج هذه الحدود هو ٠١. (١%) .

(ب) مستوى المعنوية ٠٥ر (٥%) : تعني أن احتمال وقوع مشاهدة ما في المدى (u + 26) هو ٩٥% واحتمال وقوعها خارج هذه الحدود هو ٥% أو بمعنى آخر هناك ٩٥% من مائة نتيجة فرق حقيقي ، ٥ حالات فقط نتيجة للصدفة .

(١) تحليل التباين في اتجاه واحد : One- Way Analysis of Variance

يهتم نموذج تحليل التباين في اتجاه واحد باختبار الفرق بين متوسطات العينة (K) بحيث تكون مفردات كل معاملة من مجموعة المعاملات مسحوبة عشوائياً .

وتتمثل المعادلة الخطية الممثلة لنموذج تحليل التباين في اتجاه واحد فيما يلي :

$$K_{ik} = u + a_k + e_{ik}$$

حيث :

المتوسط العام لكل مجموعات المعاملة (k) $= U$

أثر المعاملة في المجموعة المعنية (K) $= a_k$

الخطأ العشوائي المرتبط بعملية المعاينة $= e_{ik}$

(أ) جــدول ملخص لتحليل التباين في حالة تساوي المجموعات Equal Groups

S . v	S S	d f	M S	F – ratio
Between treatment groups (A)	SSA = $\sum_{K=1}^{K} \frac{Tk2}{Nk} - \frac{T2}{N}$	K - 1	MSA = $\frac{SSA}{K - 1}$	F = $\frac{MSA}{MSE}$
sampling error E	SSE = SST - SSA	N - K	MSE = $\frac{SSE}{N - K}$	
Total T	SST = $\sum_{i=1}^{N} \sum_{K=11}^{k} X2 - \frac{T2}{N}$	N - 1		

مثال :

تم توزيع ١٥ عامل فني عشوائياً في برنامج تدريبي فني على ثــلاث طرق تعليمية ، وكانت درجات اختبار الأداء للمتدربين على النحو التالي :

Instructional method	Test Scores					Total Test scores	Mean Test scores
A 1	86	79	81	70	84	400	80
A 2	90	76	88	82	89	425	85
A 3	82	68	73	71	81	375	75

HO : $U_1 = U_2 = U_3$ H1 : The Means are Not all matually equal

1- The overall mean of all 15 segments:

$$\overline{XT} = \frac{\sum X}{n} = \frac{1200}{15} = \boxed{80}$$

2- The Standard error of the mean is :

$$Sx = \sqrt{\frac{\sum (X - XT)^2}{n0 . means - 1}}$$

$$= \sqrt{\frac{(80 - 80)^2 + (85 - 80)^2 + (75 - 80)^2}{3 - 1}}$$

$$= \sqrt{\frac{50}{2}} = \boxed{5.0}$$

3- $MSB = NSX^2 = 5 (5.0)^2 = \boxed{125}$

4- From the General Formula :

$$S^2 = \frac{\sum (X - \overline{X})^2}{n - 1}$$

ويمكن حساب التباين لكل من العينات الثلاثة على النحو التالي :

$$S^2_2 = \frac{(90 - 85)2 + (76 - 85)2 + 88 - 85)2 + (82 - 85)2 + (89-85)2}{5 - 1}$$

$$= \frac{140}{4} = \boxed{35.0}$$

$$S^2_3 = \frac{(82 - 75)^2 + (68 - 75)^2 + (73 - 75)^2 + (71 - 75)^2 + (81 - 75)^2}{5 - 1}$$

$$= \frac{154}{4} \quad \boxed{38.5}$$

$$(Pould) = \frac{(n1-1)\ S^2_1 + (n^2 - s2^2) + (n3-1)\ s^23}{n1 + n2 + n3 - 3}$$

$$= \frac{(4)\ (38.5) + 4\ (35.0) + 4\ (38.5)}{5 + 5 + 5 - 3}$$

$$= \frac{448}{12} = 37.3$$

$$MSW = 37.3$$

$$F = \frac{MSB}{MSW} = \frac{125}{37.3} = 3.35$$

وحيث أن

$$F.O5\ (2,12) = 3.88$$

وعلى ذلك لا يمكن رفض النظرية الفرضية .

ولتطبيق المعاملات الواردة بالجدول ، فإنه يلزم توافر التقديرات التالية :

$n_1 = 5$ $n_2 = 5$ $n_3 = 5$ $N = 15$

$T_1 = 400$ $T_2 = 425$ $T_3 = 375$ $T = 1200$

$T^2 = 160000$ $T^2_2 = 180625$ $T^2_3 = 140625$ $T2 = 1440000$

$$\frac{T^2}{N} = \frac{1440000}{15} = 96000$$

$$\sum_{I=1}^{n} \sum_{K=1}^{k} X^2 = (86)^2 + (79)^2 + \dots\dots\dots\dots + (81)^2 = 96698$$

$$SST = \sum_{i=1}^{n} \sum_{K=1}^{k} - X^2 \frac{T^2}{N}\ 96698 - 96000 = 698$$

$$SSA = \sum_{K=1}^{k} \frac{T_K}{n_k} - \frac{T^2}{N} = \frac{160000}{5} + \frac{180625}{5} + \frac{140625}{5}$$

$$- 96000 = \boxed{250}$$

$$SSE = SST - SSA = 698 - 250 = \boxed{448}$$

ويمكن تصوير جدول تحليل التباين على النحو التالي :

S.V	S S	D f	M S	F - ratio
Between Treatmentg roups (A)	250	3 - 1 = 2	$\frac{250}{2} = 125$	
Sampling error E	448	15 – 3 = 12	$\frac{448}{12} = 37.33$	$\frac{125}{37.33} = 3.35$
Total T	698	15 – 1 = 14		

ب – **اختبار تحليل التباين في اتجاه واحد في حالة المجموعات غير المتساوية:**

أحياناً ما تكون البيانات المتاحة لتحليل التباين في اتجاه واحد غير مشتملة على أحجام متساوية من المجموعات على مستوى المعاملات المختلفة . ولتوضيح طريقة التقدير في هذه الحالة نسوق المثال التالي :

مثال : الجدول التالي يوضح متوسط عدد الكلمات المكتوبة في الدقيقة بواسطة ثلاث أنواع من الآلات الكاتبة الكهربائية خلال مدة اختبار قدرها ١٥ دقيقة أجرى على مجموعة من الأشخاص تم توزيعهم عشوائياً بعد حصولهم على دورة تدريبية ، دون أن يكون لهم سابق خبرة عن هذه الآلات .

والمطلوب اختبار مدى صحة النظرية الفرضية القائلة بأنه لا يوجد فرق بين متوسطات الكلمات المكتوبة في الدقيقة بواسطة الآلات الثلاث .

الحـــل

Typewriter rand	Average words per Minute					Total	Mean
A1	79	83	62	51	77	352	70.4
A 2	74	85	72			231	77.0
A 3	81	65	79	55		280	70.0

$n_1 = 5$ $n_2 = 3$ $n_3 = 4$ $N = 12$

$T_1 = 352$ $T_2 = 231$ $T_3 = 280$ $T = 863$

$T^2_1 = 123904$ $T^2_2 = 53361$ $T^2_3 = 78400$ $T^2 = 744769$

$$\frac{T^2}{N} = \frac{7.44769}{12} \boxed{= 62064.1}$$

$$\sum_{i=1}^{n} \sum_{K=1}^{k} X^2 = 79^2 + 83^2 + \ldots\ldots + 55^2 = \boxed{63441}$$

$$SST = \sum_{i=1}^{n} \sum_{K=1}^{k} X2 - \frac{T^2}{N}$$

$$= 63441 - 62064.1 = \boxed{1376.9}$$

$$SSA = \sum_{K=1}^{k} \frac{T^2{}_K}{n_k} - \frac{T^2}{N}$$

$$= \frac{123904}{5} + \frac{53361}{3} + \frac{78400}{4} - 62064.1 = 103.7$$

$$SSE = SST - SSA$$
$$SSE = 1376.9 - 103.7 = \boxed{1273.2}$$

وعلى ذلك يمكن تصوير جدول تحليل التباين على النحو التالي :

Anova table

S V	S S	D f	M S	F - ratio
A E	103.7 1273.2	$3 - 1 = 2$ $12 - 3 = 9$	$\frac{103.7}{2} = 51.8$ $\frac{1273.2}{9} = 141.5$	$\frac{51.8}{141.5} = 0.37$
Total	1376.9	$12-1 = 11$		

الفصل الخامس عشر

تطبيقات عملية

أولاً : الأسئلة النظرية :

١- ما هو مفهوم علم الإحصاء ؟

٢- تحدث عن العلاقة بين علم الإحصاء والعلوم الأخرى ؟

٣- ناقش بإيجاز طرق جمع البيانات الإحصائية ؟

٤- إشرح أنواع البيانات الإحصائية ؟

٥- ما هي الأنواع المختلفة للعينات ؟

٦- تحدث عن الأخطاء الشائعة في جمع البيانات ؟

٧- ما هي أهم مصادر الحصول على البيانات الإحصائية ؟

٨- إشرح بإيجاز أهم طرق عرض البيانات الإحصائية ؟

٩- ما هي أنواع الاستمارات الإحصائية ؟

١٠- ما الفرق بين المتغير المتصل والمنفصل ؟

١١- ما هي أهم مستويات قياس المتغيرات ؟

١٢- ما الفرق بين المقياس الاسمي والمقياس الفتري ؟

١٣- ما الفرق بين المتغير المستقل والمتغير التابع ؟

١٤- عرف المتغير وفرق بينه وبين الثابت ؟

١٥- ما هي الأقسام المختلفة لعلم الإحصاء ؟

١٦- ما الفرق بين العينة العشوائية والعينة الطبقية ؟

١٧- أذكر فوائد التوزيعات التكرارية ؟

١٨- عرف الأرقام القياسية ؟

١٩- فرق بين المدرج والمضلع والمنحنى التكراري ؟

٢٠- عرف الوسط والوسيط والمنوال ؟ وما العلاقة بينهم ؟

٢١- ما المقصود بالتشتت وما هي أهم المقـاييس الإحصـائية المسـتخدمة لقياسه ؟

٢٢- متى تستخدم سلسلة الأرقام القياسية ؟

٢٣- ما هي أهداف استخدام مقاييس التشتت ؟

٢٤- اشرح بإيجاز الأشكال المختلفة للأرقام القياسية ؟

٢٥- ما معنى معامل الالتواء ؟

٢٦- ما المقصود بمجتمع العينة ؟

٢٧- عرف الارتباط وأذكر الأشكال المختلفة للانتشار ؟

٢٨- ما الفرق بين التعداد النظري والتعداد الفعلي ؟

٢٩- أذكر بإيجاز الخطوات الأساسية في البحث الإحصائي ؟

٣٠- ما معنى تحليل البيانات وتفسيرها ؟

٣١- ما المقصود بالإحصاء الوصفي والإحصاء الرياضي ؟

٣٢- ما هي الخصائص الأساسية للمنوال ؟

٣٣- أذكر خصائص المتوسط الهندسي ؟

٣٤- وضح خصائص الانحراف الربيعي ؟

٣٥- وضح خصائص الانحراف المعياري ؟

٣٦- ما هي أوجه الاختلاف بين الانحراف الربيعي والانحراف المعياري؟

٣٧- ما المقصود بالاحتمالات وما الفرق بين الأحداث المتنافية والمستقلة ؟

ثانياً : تمارين محلولة

مثال (١) :

من الجدول التالي أحسب معادلة خط الاتجاه المستقيم لكمية الصادرات لإحدى الدول بالألف طن في الفترة من سنة ٢٠٠٠ إلى سنة ٢٠٠٤ وذلك باستخدام طريقة المربعات الصغرى .

السنة	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤
كمية الصادرات بالألف طن	٢٠	٣٠	٤٠	٥٠	٦٠

الإجابة :

السنة	ص	س	س ص	س٢
٢٠٠٠	٢٠	٠	٠	٠
٢٠٠١	٣٠	١	٣٠	١
٢٠٠٢	٤٠	٢	٨٠	٤
٢٠٠٣	٥٠	٣	١٥٠	٩
٢٠٠٤	٦٠	٤	٢٤٠	١٦
	٢٠٠	١٠	٥٠٠	٣٠

ص = أ + ب س   معادلة الخط المستقيم

مجـ ص = ن أ + ب مجـ س   (١)

مجـ س ص = أ مجـ س + ب مجـ س٢   (٢)

بالتعويض في المعادلة (١) ، (٢)

٢٠٠ = ٥ × أ + ب × ١٠   (٣)

٥٠٠ = ١٠ أ + ٣٠ ب   (٤)

بضرب المعادلة ٣ في ٢ وبالطرح ينتج :

٤٠٠ = ١٠ أ + ٢٠ ب

١٠٠ = ١٠ ب

إذا ب = ١٠

بالتعويض في المعادلة ٣ بقيمة (ب) :

٢٠٠ = ٥ أ + ١٠٠

١٠٠ = ٥ أ

إذاً أ = ٢٠

بالتعويض عن أ ، ب بقيمها في معادلة الخط المستقيم

إذاً ص = ٢٠ + ١٠ س

نقطة الأصل سنة ٢٠٠٠ والوحدة الزمنية سنة ، ويمكن استخدام نقطة الأصل في الوسط كما يلي :

السنة	ص	س	س ص	س٢
٢٠٠٠	٢٠	– ٢	– ٤٠	٤
٢٠٠١	٣٠	– ١	– ٣٠	١
٢٠٠٢	٤٠	صفر	صفر	٠
٢٠٠٣	٥٠	١	٥٠	١
٢٠٠٤	٦٠	٢	١٢٠	٤
	٢٠٠	صفر	– ٧٠٠ + ١٧٠ ١٠٠	١٠

ص = أ + ب س   معادلة الخط المستقيم

مجـــ ص = ن أ + ب مجـــ س   (٢)

مجـــ س ص = أ مجـــ س + ب مجـــ س٢   (٣)

بالتعويض في المعادلة ١ ، ٢

٢٠٠ = ٥ أ   إذا أ = ٤٠

١٠٠ = ١٠ ب   إذا ب = ١٠

بالتعويض في معادلة الخط المستقيم

إذاً ص = ٤٠ + ١٠ س

نقطة الأصل سنة ٢٠٠٢ والوحدة الزمنية سنة

ويمكن تحول هذه المعادلة إلي الصورة السابقة كما يلي :

- معامل س و هي ١٠ لا يتغير في الحالتين لأن الوحدة الزمنية واحدة .
- نقطة الأصل تغيرت من ٢٠٠٠ إلى ٢٠٠٢ فيكون الفرق ٢

إذاً ص = ٤٠ + (١٠ × - ٢) + ١٠ س

إذاً ص = ٤٠ + (- ٢٠) + ١٠ س

= ٢٠ + ١٠ س

أي أن معادلة خط الاتجاه الجديدة على أساس أن نقطة الأصل سنة ٢٠٠٠ والوحدة الزمنية سنة ... هـــي :

ص = ٢٠ + ١٠ س

نقطة الأصل سنة ٢٠٠٠ والوحدة الزمنية سنة

وهي نفس الصورة السابقة .

مثال (٢) : فيما يلي أجور عدد من العمال في أحد المصانع ، والمطلوب إيجاد الوسط الحسابي للأجور بطريقتين مختلفتين :

فئات الأجور	٥-١٠	١٠-١٥	١٥-٢٠	٢٠-٢٥	٢٥-٣٠
عدد العمال	٥٢	٥٦	٥٩	٦١	٤٧

فئات الأجور	٣٠-٣٥	٣٥-٤٠	٤٠-٤٥
عدد العمال	١٢	٧	٥

الحـــل

أولاً : الحل باستخدام الطريقة المباشرة :

ف	ك	س	س × ك
٥ -	٥٢	٧,٥	٣٩٠
١٠ -	٥٦	١٢,٥	٧٠٠
١٥ -	٥٩	١٧,٥	١٠٣٢,٥
٢٠ -	٦١	٢٢,٥	١٣٧٢,٥
٢٥ -	٤٧	٢٧,٥	١٢٩٢,٥
٣٠ -	١٢	٣٢,٥	٣٩٠
٣٥ -	٧	٣٧,٥	٢٦٢,٥
٤٠ - ٤٥	٥	٤٢,٥	٢١٢,٥
	٢٩٩		٥٦٥٢,٥

$$\bar{س} = \frac{\text{مجـ (س × ك)}}{\text{مجـ ك}}$$

$$\frac{٥٦٥٢,٥}{٢٩٩} = ١٨,١٩$$

ثانياً : الحل باستخدام الطريقة المختصرة :

ف	ك	س	ح	ح × ك
٥ –	٥٢	٧,٥	– ١٥	– ٧٨٠
١٠ –	٥٦	١٢,٥	– ١٠	– ٥٦٠
١٥ –	٥٩	١٧,٥	– ٥	– ٢٩٥
٢٠ –	٦١	٢٢,٥	صفر	صفر
٢٥ –	٤٧	٢٧,٥	٥	٢٣٥
٣٠ –	١٢	٣٢,٥	١٠	١٢٠
٣٥ –	٧	٣٧,٥	١٥	١٠٥
٤٠–٤٥	٥	٤٢,٥	٢٠	١٠٠
	٢٩٩			– ١٠٧٥

$$\text{س} = \text{١} \; \frac{\text{مجـ (ح × ك)}}{\text{مجـ ك}}$$

$$= \text{٢٢,٥} + \frac{\text{– ١٠٧٥}}{\text{٢٩٩}}$$

$$= \text{٢٢,٥} - \text{٣,٥٩}$$

$$= \text{١٨,٩١}$$

مثال(٣) : من البيانات التالية أوجد قيمة الوسيط بطريقتين مختلفتين

ف ٢–   ٤–   ٧–   ١٢–   ١٧–   ٢٢ وأقل من ٣٢

ك ٦٠   ١٢٠   ١٥٠   ١٣٠   ١١٠   ٨٠

الحـــل

يتم حل هذا التمرين باستخدام التكــرار المتجمــع الصــاعد والهــابط وللسهولة سوف نضع التكرار المتجمع الصاعد والتكرار المتجمع الهابط فــي جدول واحد كما هو واضح فيما يلي

ف		متجمع صاعد		متجمع هابط	
		الحدود العليا للفئات	التكرار المتجمع الصاعد	الحدود السفلي للفئات	التكرار المتجمع الهابط
٢ –	٦٠	أقل من ٤	٦٠	٢ فأكثر	٦٥٠
٤ –	١٢٠	أقل من ٧	١٨٠	٤ فأكثر	٥٩٠
٧ –	١٥٠	أقل من ١٢	٣٣٠	١٠فأكثر	٤٧٠
١٢–	١٣٠	أقل من ١٧	٤٦٠	١٢ فأكثر	٣٢٠
١٧–	١١٠	أقل من ٢٢	٥٧٠	١٧فأكثر	١٩٠
٢٢وأقل من ٣٢	٨٠	أقل من ٣٢	٦٥٠	٢٢فأكثر	٨٠
	٦٥٠				

ترتيب الوسيط = $\frac{\text{مجـ ك}}{٢} = \frac{٦٥٠}{٢} = ٣٢٥$

قيمة الوسيط : **أولا : بإستخدام التكرار المتجمع الصاعد :**

الوسيط = $٧ + \frac{٣٢٥ - ١٨٠}{٣٣٠ - ١٨٠} \times ٥$

$$= ٧ + \frac{١٤٥}{١٥٠} \times ٥ = ٧ + ٤,٨٣ = ١١,٨٣$$

ثانيا : باستخدام التكرار المتجمع الهابط

$$= ٧ + \frac{٤٧٠ - ٣٢٠}{٤٧٠ - ٣٢٥} \times ٥$$

$$= ٧ + \frac{١٤٥}{١٥٠} \times ٥ = ٧ + ٤,٨٣ = ١١,٨٣$$

مثال (٤) : من البيانات التالية أوجد قيمة كل من الربيع الأدنى والربيع الأعلى ونصف المدى الربيعي .

ف	٨-	١٥-	٢٥-	٣٥-	٤٥-	٦٥-	١٠٠
ك	٢٨	٤٢	٢٥	١٢	٧	٤	٧

الحـــل

ف	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد
٨-	٢٨	أقل من ١٠	٢٨
١٠-	٤٢	أقل من ١٥	٧٠
١٥-	٢٥	أقل من ٢٥	٩٥
٢٥-	١٢	أقل من ٣٥	١٠٧
٣٥-	٧	أقل من ٤٥	١١٤
٤٥-	٤	أقل من ٦٥	١١٨
٦٥-١٠٠	٢	أقل من ١٠٠	١٢٠
	١٢٠		

ترتيب الربيع الادني = $\frac{١٢٠}{٤}$ = ٣٠

ترتيب الربيع الأعلى = $\frac{١٢٠ \times ٣}{٤}$ = ٩٠

قيمة الربيع الأدنى = ١٠+ $\frac{٣٠ - ٢٨}{٧٠ - ٢٨}$ × ٥

= ١٠ + ٢٤ر = ١٠,٢٤

قيمة الربيع الأعلى = ١٥+ $\frac{٩٠ - ٧٠}{٩٥ - ٧٠}$ ×١٠

= ١٥+٨= ٢٣

نصف المدى الربيعي = $\frac{٢٣-١٠,٢٤}{٢}$ = $\frac{١٢,٧٦}{٢}$ ٦,٣٨

مثال (٥) : الجدول الآتي يبين الأجر الشهري بالجنيهات لمجموعة من العمال في أحد المصانع والمطلوب إيجاد الانحراف المعياري للأجور بالمصنع :

الأجر بالجنيه	٨-	١٠-	١٥-	١٨-	٢٠-	٢٥-	٣٠-٥٠
عدد العمال	٧	١٥	٣٢	٤٣	٢٨	١٢	٣

الحـــل

ف	ك	س	ح	ح × ك	ح٢ × ك
٨ –	٧	٩	– ١٠	– ٧٠	٧٠٠
١٠ –	١٥	١٢,٥	– ٦,٥	– ٩٧,٥	٦٣٢,٧٥
١٥ –	٣٢	١٥,٥	– ٣,٥	– ٨٠	٢٠٠
١٨ –	٤٣	١٩	صفر	صفر	صفر
٢٠ –	٢٨	٢٢,٥	٣,٥	٥٨	٣٤٣
٢٥ –	١٢	٢٧,٥	٨,٥	١٠٢	٨٦٧
٣٠ وأقل من ٥٠	٣	٤٠	٢١	٦٣	١٣٢٣
	١٤٠			١٥,٥	٤٠٦٦,٧٥

$$ع = \sqrt{\frac{٤٠٦٦,٧٥}{١٥٠} - \left\{\frac{١٥,٥}{١٥٠}\right\}^{٢}}$$

$$\sqrt{٢٧,١١٧ - .٠١} = ٥,٢$$

مثال (٦) : من الجدول التكراري التالي أحسب معامل التشتت :

فئـــات	أقل من ١٠	١٠-	٢٠-	٣٠-	٤٥-	٦٠-	٩٠ فأكثر
تكرارات	٧	٩	١٢	١٥	٢٢	١٠	٣

الحـــل

جدول تكراري متجمع صاعد

فئات	تكرارات	حدود عليا للفئات	تكرار مجتمع صاعد
أقل من ١٠	٧	أقل من ١٠	٧
١٠-	٩	أقل من ٢٠	١٦
٢٠-	١٢	أقل من ٣٠	٢٨
٣٠-	١٥	أقل من ٤٥	٤٣
٤٥-	٢٢	أقل من ٦٠	٦٥
٦٠-	١٠	أقل من ٩٠	٧٥
٩٠ فأكثر	٣	أقل من الحد الأعلى	٧٨
	٧٨		

ترتيب الربيع الأول الأدنى = $\frac{\text{مجـ ك}}{٤}$ = $\frac{٧٨}{٤}$ = ١٩,٥

قيمة الربيع الأول = ٢٠ + (١٩,٥ − ١٦) / (٢٨ − ١٦) × ١٠

ر١ = ٢٠ + ٣,٥ / ١٢ × ١٠ = ٢,٩

ر = ٢٢,٩

ترتيب الربيع الثالث (الأعلى) = (مجـ ك × ٣) / ٤ = (٧٨ × ٣) / ٤

= ٥٨,٥

قيمة الربيع الثالث (الأعلى) = ٤٥ + (٥٨,٥ − ٤٣) / (٦٥ − ٤٣) × ١٥

ر٣ = ٤٥ + ١٥,٥ / ٢٢ × ١٥ = ٤٥ + ١٠,٥٦٨

ر٣ = ٥٥,٥٦٨

$$\text{معامل التشتت الربيعي} = \frac{٣ - ١}{٣ + ١} \times ١٠٠$$

$$= \frac{٢٢,٩ - ٥٥,٥٦٨}{٢٢,٩ + ٥٥,٥٦٨} \times ١٠٠$$

$$= \frac{٣٢,٦٦٨}{٧٨.٤٦٨} \times ١٠٠ = ٤١,٦٣\%$$

مثال (٧) : أوجد معامل ارتباط الرتب لسبيرمان من البيانات التالية والتي تمثل درجات خمسة طلاب من شعبة المحاسبة بكلية التجارة في مـادتي المحاسـبة الحكومية والمالية :

محاسبة حكومية	٧٤	٨٩	٧٥	٨٢	٩١
محاسبة مالية	٧٧	٨٣	٧٤	٩٠	٨٤

الحـــــل

محاسبة حكومية	محاسبة مالية	رتب س	رتب ص	الفرق ف	ف٢
٧٤	٧٧	٥	٤	١	١
٨٩	٨٣	٢	٣	١	١
٧٥	٧٤	٤	٥	١	١
٨٣	٩٠	٣	١	٢	٤
٩١	٨٤	١	٢	١	١
					٨

معامل ارتباط الرتب (ر) = $١ - \frac{٦ \text{ مجـ } ف^{٢}}{ن(ن^{٢} - ١)}$ (حيث ن عدد المفردات)

$ر = ١ - \frac{٦ \times ٨}{٥ \times ٢٤} = ٦ر$

وهذه النتيجة توضح أن هناك علاقة طردية بين الظاهرتين .

مثال (٨) : فيما يلي تقديرات عشر طلاب في مـــادتي الإحصـــاء والمحاســـبة والمطلوب إيجاد معامل ارتباط الرتب بين تقديرات المادتين :

تقدير	**جيد جداً**	**ممتاز**	**مقبول**	**جيد جداً**	**ضعيف جداً**	**ضعيف**	**مقبول**
تقدير	ممتاز	جيد	جيد جداً	مقبول	ضعيف	مقبول	مقبول
	جيد	جيد	مقبول				
	ضعيف	مقبول	ضعيف جداً				

الحـــل

تقدير الإحصاء س	تقدير المحاسبة ص	رتب س	رتب ص	الفرق بينهما	ف٢
جيد جداً	ممتاز	٢,٥	١	١,٥	٢,٢٥
ممتاز	جيد	١,٠	٣	٢	٤
مقبول	جيد جداً	٧	٢	٥	٢٥
جيد جداً	مقبول	٢,٥	٥,٥	٣	٩
ضعيف جداً	ضعيف	١٠	٨,٥	١,٥	٢,٢٥
ضعيف	مقبول	٩	٥,٥	٣,٥	١٢,٢٥
مقبول	مقبول	٧	٥,٥	١,٥	٢,٢٥
جيد	ضعيف	٤,٥	٨,٥	٤	١٦
جيد	مقبول	٤,٥	٥,٥	١	١
مقبول	ضعيف جداً	٧	١٠	٣	٩
					٨٣

ملحوظـــــة :

كيفية إيجاد الرتب هي إذا تساوي تقديران نجمع رتبيهما ونقسم الناتج على ٢ .

وإذا تساوت ثلاث تقديرات تجمع رتبها ونقسمها على ثلاث :

$$\text{ر} = ١ - \frac{\text{٦ مجـ ف}^{٢}}{\text{ن} (\text{ن}^{٢} - ١)}$$

$$= ١ - \frac{٦ \times ٨٣}{١٠ (١٠٠ - ١)} = ١ - \frac{٦ \times ٨٣}{١٠ \times ٩٩}$$

= ١ – ٥٠٣ر = ٤٩٧ر

مثال (٩) : فيما يلي أسعار وكميات ثلاث سلع أ ، ب ، ج عامي ٢٠٠-٢٠٠٤

السنة	سلعة أ		سلعة ب		سلعة ج	
	السعر	الكمية	السعر	الكمية	السعر	الكمية
٢٠٠٠	١	٢	٣	٥	١	٤
٢٠٠٤	٣	٣	٣	٦	٣	٥

والمطلـــوب :

(أ) حساب الرقم القياسي التجميعي للأسعار المرجح بكميات سنة الأســـاس (رقم لاسبير) .

(ب) حساب الرقم القياسي التجميعي للأسعار والمرجح بكميات سنة المقارنة (رقم باش) .

(ج) حساب الرقم القياسي الأمثل (فيشر)

الحـــل

	سلعة أ		سلعة ب		سلعة ج		
	السعر ع	الكمية ك	السعر ع	الكمية ك	السعر ع	الكمية ك	
سنة ٢٠٠٠	١	٣	٢	٥	١	٤	المجموع
سنة ٢٠٠٤	٢	٣	٣	٦	٣	٥	
ع ك	٦		١٦		١٢		٣٣
ع.ك.	٣		١٠		٤		١٧
ع ك	٦		١٨		١٥		٣٩
ع.ك.	٣		١٢		٥		٢٠

أ - الرقم التجميعي للأسعار المرجح بكميات سنة الأساس (سبير) .

$$= \frac{\text{مجـ ع ك}}{\text{مجـ ع. ك .}} \times ١٠٠$$

$$\frac{٣٣}{١٧} \times ١٠٠ = ١٩٤,١٢\%$$

أ = رقم لاسبير $= \frac{\text{مجـ ع ك .}}{\text{مجـ ع. ك.}} \times ١٠٠ = \frac{٦٢}{\text{...}} \times ١٠٠ = ١٣١,٩\%$

ب – رقم باش $= \frac{\text{مجـ ع ك}}{\text{مجـ ع. ك.}} \times ١٠٠ = \frac{٨٨}{٦٧} \times ١٠٠ = ١٣١,٣\%$

ج – الرقم القياسي الأمثل لفيشر .

$$= \sqrt{\frac{\text{مجـ ع ك .}}{\text{مجـ ع . ك .}} \times \frac{\text{مجـ ع ك}}{\text{مجـ ع . ك}}} \times ١٠٠$$

$$= \sqrt{\frac{٦٢}{٤٧} \times \frac{٨٨}{٦٧}} \times ١٠٠ = \sqrt{\frac{٥٤٥٦}{٢١٤٩}} \times ١٠٠$$

$$= \sqrt{١,٧٣٢٦} \times ١٠٠ = ١,٣١٦٢ \times ١٠٠$$

$= ١٣١,٦٣\%$

مثال (١١) : يمثل الجدول التالي مبيعات شركة خالد التجارية في الفترة من ٩٧ إلى ٢٠٠٤ :

السنة	٩٧	٩٨	٩٩	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤
المبيعات	١٠	١٥	٢٥	٥٠	٧٠	١١٠	١٢٠	١٤٠

		٢٠٠٠	٢٠٠٤
سلعة (أ)	السعر	٣	٤
	الكمية	٢	٣
سلعة (ب)	السعر	٦	٨
	الكمية	٤	٥
سلعة (ج)	السعر	٥	٦
	الكمية	١	٢
سلعة (د)	السعر	٣	٤
	الكمية	٤	٦

الحــــــل

السنة ٢٠٠٠	٢٠٠٤	ع ك	ع .ك.	ع ك	ع.ك.
السعر ٣	٤				
الكمية ٢	٣	٨	٦	١٢	٩
السعر ٦	٨				
الكمية ٤	٥	٣٢	٢٤	٤٠	٣٠
السعر ٥	٦				
الكمية ١	٢	٦	٥	١٢	١٠
السعر ٣	٤				
الكمية ٤	٦	١٦	١٢	٢٤	١٨
		٦٢	٤٧	٨٨	٦٧

ب – الرقم التجميعي للأسعار المرجح بكميات سنة المقارنة (باش) .

$$= \frac{\text{مجـ ع ك}}{\text{مجـ ع . ك}} \times ١٠٠$$

$$= \frac{٣٩}{٢٠} \times ١٠٠ = ١٩٥\%$$

ج – الرقم القياسي الأمثل لفيشر =

$$\sqrt{\frac{\text{مجـ ع ك .}}{\text{مجـ ع . ك .}} \times \frac{\text{مجـ ع ك .}}{\text{مجـ ع . ك}} \times ١٠٠}$$

$$= \frac{٣٣}{١٧} \times \frac{٣٩}{٢٠} = ١٠٠$$

$$= \frac{١٢٨٧}{٣٤٠} \times ١٠٠ = ١,٩٤٥٥٨ \times ١٠٠ = ١٩٤,٥٥٨ \%$$

مثال (١٠) : فيما يلي أسعار كميات أربع سلع أ ، ب ، ج ، د والمطلوب :

(أ) حساب الرقم القياسي للأسعار المرجح بكميات سنة الأساس (لاسبير) .

(ب) حساب الرقم القياسي للأسعار المرجح بكميات سنة المقارنة (باش) .

(ج) حساب الرقم الثابت الأمثل لفيشر .

والمطلوب : إيجاد معادلة خط الاتجاه العام بطريقة المربعات الصغرى .

الحـــل

السنة	المبيعات	س	س ص	س٢	ص٢	أثر الاتجاه العام
٩٧	١٠	٧-	٧٠-	٤٩	٣,٧٦-	١٣,٧٦+
٩٨	١٥	٥-	٧٥-	٢٥	١٦,٦	١,٦-
٩٩	٢٥	٣-	٧٥-	٩	٣٦,٩٦	٩,٩٦ -
٢٠٠٠	٥٠	١-	٥٠-	١	٥٧,٣٢	٧,٣٢-
٢٠٠١	٧٠	١	٧٠	١	٧٧,٦٨	٧,٦٨
٢٠٠٢	١١٠	٣	٣٣٠	٩	٩٨,٠٤	١١,٩٦+
٢٠٠٣	١٢٠	٥	٦٠٠	٢٥	١١٨,٤	١,٦+
٢٠٠٤	١٤	٧	٩٨٠	٤٩	١٣٨,٧٦	١,٢٣+
المجموع	٥٤٠	صفر	١٧١٠	١٦٨	٥٤٠	

^ص = م س + جـ

م = (مجـ س ص) / (مجـ س ٢) = ١٧١٠ / ١٦٨ = ١٠,١٨

ج = (مجـ ص) / ٨ = ٥٤٠ / ٨ = ٦٧,٥

^ص = ١٠,١٨ س + ٦٧,٥

ثالثاً : تمارين عملية غير محلولة :

تمارين على التوزيعات التكرارية

١-تمثل البيانات التالية درجات ٤٠ طالباً في مادة الإحصاء والمطلوب إنشاء توزيع تكراري لهذه البيانات :

٦١	٧٨	٥٩	٦١	٧٦	٦٧	٥٥	٣٩
٥٥	٦٣	٩٩	٦٩	٨٢	٦٣	٨٤	٥٦
٤٩	٦٥	٥٠	٥٠	٧٠	٧٧	٥٩	٥٩
٢١	٢٤	٤٨	٤٧	٤٢	٤٢	٢١	٥٠
٥٧	٦٧	٣٣	٣٣	٥٧	٢٤	٢١	٥٢

٢-يبين التوزيع التكراري التالي للأجور الأسبوعية بالجنيه لعدد من العاملين في شركة عمر الكبرى :

فئات الأجور	أقل من ٥٠	٦٠-	٧٠-	٨٠-	٩٠-	١٠٠-	١٠٠-١١٠
عدد العاملين	٤٠	٥٩	٨٥	١٤٥	٥٥	٣٠	٦

والمطلـــــوب :

(أ) رسم المدرج التكراري .

(ب) رسم المضلع التكراري .

(ج) رسم المنحنى التكراري

(د) رسم المنحي المتجمع الصاعد والمتجمع الهابط .

(ه) تحديد عدد العاملين الذين يحصلون على أجر أسبوعي قدره ٩٠ جنيهاً فأكثر .

٣- يبين التوزيع التكراري التالي فئات الأعمار لعدد من السكان في منطقة معينة .

فئات الأعمال	٢٠-	٣٠-	٤٠-	٥٠-	٦٠-	٧٠-٨٠
عدد السكان	٩٠	١٨٠	١٣٥	١٢٠	٧٠	٢٠

والمطلوب معرفة :

(أ) النسبة المئوية للسكان الذين تقل أعمارهم عن ٤٠ سنة .

(ب) النسبة المئوية للسكان الذين يكون أعمارهم ٦٠ سنة فأكثر .

٤- يمثل الجدول التكراري الآتـــي المصاريف العمومية بالجنيهات لعدد ٧٤ شركة في سنة ٢٠٠٣

المصاريف العمومية	عدد الشركات
١-	٣
١٥-	٥
٢٠-	٨
٢٥-	١٢
٣٠-	١٩
٣٥-	١٠
٤٠-	٨
٤٥-	٦
٥٠-	٢
٥٥-	١

والمطلوب :

١- رسم المدرج التكراري والمضلع التكراري .

٢- إنشاء جدول تكراري متجمع صاعد وجدول تكراري متجمع هابط .

٣- رسم المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط .

(٥) : فيما يلي بيان بإنتاج شركة العباسي للسلع الكهربائية بالآلف وحدة :

السنة	١٩٩٥	١٩٩٦	١٩٩٧	١٩٩٨	١٩٩٩	٢٠٠٠	٢٠٠١	٢٠٠٢
الإنتاج	٥	١٠	١٠	٢٠	٥٠	٦٠	٧٥	٩٠

والمطلوب عرض البيانات السابقة باستخدام :

أ - الخط البيانـــــي   ب - الأعمدة البسيطة

(٦) : الجدول التالي يبين الأهمية النسبية لبنود الاستثمار فـــي بنـــك فريـــد التجاري الدولي ، أذون خزانة ٢٠% وصكوك حكومية ٣٠% وصكوك أخرى ١٠% وسندات مؤسسات عامة ٤٠% .

والمطلوب رسم دائرة تعكس مساحتها الأهمية النسبية لمكونات الاستثمار فـــي البنك .

(٧) : الجدول التالي يبين توزيع الطلبة بإحدى الكليات حسب النوع والفرقـــة الدراسية في عام ٢٠٠٠ والمطلوب توضيح هذه البيانـــات بيانيـــاً بطريقـــة مناسبة توضح المجموع .

الفرقة / النوع	طلبـــة	طالبات	المجموع
الأولى	٢٥٠	٢٠٠	٤٥٠
الثانية	٢٠٠	١٨٠	٣٨٠
الثالثة	١٧٥	١٢٠	٢٩٥
الرابعة	١٣٠	٩٠	٢٢٠

(٨) : الجدول التالي يبين حجم الاستثمارات في شركة النصر التجارية عامي ٢٠٠٠ و ٢٠٠٤ (بالألف) والمطلوب تمثيل ذلك بدائرتين تخص كل سنة دائرة ؟

بيان	معاملات محلية	معاملات خارجية	معاملات مشتركة
١٩٩٨	٤٠٠	٣٠	٧٠
٢٠٠٤	٥٠٠	١٢٠	١٨٠

تمارين على مقاييس النزعة المركزية :

١ - أحسب الوسط الحسابي والانحراف المعياري للتوزيع التالي :

فئات	تكرار
٣-	٧
٦-	٩
٩	١٢
١٢	٢٥
١٥	١٣
١٨-٢١	٤

إحسب أيضاً معامل الاختلاف بطريقتين مختلفتين وكذلك المنوال بطريقتين مختلفتين .

٢ - أحسب الوسط الحسابي والانحراف المعياري ومعامل الاختلاف للتوزيـــع التكراري التالي :ـــ

فئات	تكرار
٠ -	٣٠
١٠ -	٤٠
٣٠ -	٢٥
٥٠ - ١٠٠	٥

٣ – أحسب الوسط الحسابي والانحراف المعياري لتوزيع عدد ساعات العمل الأسبوعية لمائة عامل ، وكذلك :

(أ) معامل الالتواء وعلق عليه .

(ب) معامل التفرطح وعلق عليه .

(ج) المنوال والوسيط بالرسم .

(د) معامل الاختلاف بطريقتين مختلفتين .

عدد الساعات	عدد العمال
١١-	٢
١٤-	١
١٧-	٢
٢٠-	١
٢٣-	٦
٢٦-	٧
٢٩-	١١
٣٢-	١٢
٣٥-	١٣
٣٨-	١٧
٤١-	١٩
٤٤-٤٧	٩

٤- قارن بين توزيعي قيم مجموعتي (أ) ، (ب) مستخدماً معامل الالتواء والتفرطح وأذكر أي المجموعتين أكثر تشتتاً وارسم المنحنى التكراري لبيان الالتواء وإشرحه .

القيمة	تكرار مجموعة أ	تكرار مجموعة ب
٥ -	٢٢	١٣
١٠ -	٣٥	١٧
١٥ -	٤٧	٢٥
٢٠ -	٥٢	٢٦
٢٥ -	٢٨	٣٠
٣٠ -	١٥	٣٢
٣٥ -	٢٠	٣٥
٤٠ -	١٥	٣٧
٤٥ -	٢٢	٣٩
٥٠ -	١٠	٥٠
٥٥ -	١٩	٣٧
٦٠ -	١١	٣٠
٦٥-٧٠	٧	٢٠

٥- أحسب المتوسط الحسابي في الجدول التكراري الآتي بأية طريقة تجدها مناسبة وأوجد معامل الاختلاف .

الفئات	التكرار
أقل من ٢٠	١٦
٢٠ –	٢٢
٢٥ –	٢١
٣٠ –	٢٥
٣٥ –	٣٥
٤٠ –	٤٢
٤٥ –	٣٢
٥٠ –	٣٠
٥٥ –	٢٢
٦٠ –	٢١
٦٥ –	٢٠
٧٠ فأكثر	١٤
المجموع	٣٠٠

٦ – فيما يلي درجات ٦٠ شخصاً في اختبار لمادة الإحصاء :

٧٥	٤٦	٣٨	٢٤	٧٠	٩٢
٣٢	٦٤	٢٩	١٧	٨٥	٤٩
٢٥	٧٥	٦٤	٢٨	٢٥	٥٥
٥٤	٨٢	٢٥	٣٥	٣٣	٤٦
٨٢	٨٤	٣٦	٦٠	٩٠	٧٢
١٥	٧٣	٤٤	٥٢	٤٥	٨٤
٢٦	٢٥	١٥	٣٤	٥٢	٥٠
٣٤	٣٦	٩٥	٥٧	٣٦	٦٤
٤٧	٤٤	٢٧	٣٢	٤٨	٧٢
٨٢	٧٦	٨٩	٦٠	٧٠	٨٣

أحسب المتوسط الحسابي لهذه الدرجات ثم فرغها في جدول تكراري وأحسب المتوسط الحسابي من هذا الجدول (أبدأ الجدول بالدرجة ١٥ متخذاً طول كل فئة ١٠ درجات) وقارن بين الناتجين .

تمارين على الارتباط والانحدار

١ – أوجد كلا من معادلتي خط انحدار :

ص على س ، س على ص من البيانات التالية :

س	١	٢	٤	٧	٩	١١	١٣	١٥
ص	١	٤	٦	٨	١٠	١٢	١٤	١٦

وكذلك أحسب معامل الارتباط بين قيم س ، ص .

٢ – أحسب معامل الارتباط (بيرسون) بين قيم س ، ص من البيانات التالية ، وأحسب معامل ارتباط سبيرمان لنفس القيم وقارن بينهما .

س	١١٣	١١٤	١١٥	١١١	١١٦	١١٢	١١٣
ص	٣١٢	٣١٦	٣١٥	٣١٢	٣١٤	٣١٥	٣١٧

٣ – فيما يلي بيان التقديرات التي حصل عليها ثمانية من الطلبة في مـــادتي الاقتصاد وإدارة الأعمال والمطلوب حساب معامل الارتباط بين تقديرات هاتين المادتين :

تقدير الاقتصاد	ض	ل	م	ج ج	ج	ل	ل	ض ج
تقدير إدارة الأعمال	ض	ل	ج ج	م	ل	ج	ض	ض ج

حيث نرمز إلى امتياز (م) ، جيد جداً (ج ج) ، جيد (ج) مقبول (ل) ضـــعيف (ض) ، ضعيف جداً (ض ج) .

٤ – قدر الصادرات عام ١٩٩٩ إذا توافرت لديك البيانات التالية :

السنـــة	٨٦	٨٧	٨٨	٨٩	٩٠	٩١	٩٢	٩٣
إجمالي الصادرات	٢٣	٢٥	٢٦	٢٩	٣٠	٣٢	٣٤	٣٦

٥- فيما يلي بيان بإجمالي المنفق من ميزانية الأسرة (س) والإنفاق على
المأكل (ص) في عينة تشمل ١٠ أسر ، أوجـــد :
(أ) معامل ارتباط الرتب (سبيرمان) .
(ب) معامل ارتباط بيرسون بالطريقة المباشرة وبطريقة الانحرافات عن أوساط فرضية .

إجمالي المنفق من الميزانية	٢٠	٢٥	٣٠	٣٥	٤٠	٤٥	٥٠	٥٥	٦٠	٦٥
الإنفاق على المآكل	١٤	١٥	١٧	١٦	١٨	٢٠	٢١	٢٥	٢٤	٣٠

٦- الجدول التالي يظهر العلاقة بين المرتب الشهري وجملـــة الاســـتقطاعات الشهرية بالجنيهات لعشرين من الموظفين .

الاستقطاعات / المرتب	٢٠-	٢٥-	٣٠-	٣٥-٤٠	المجموع
٩٠-	٢	٣			٥
١١٠-		٢	١	١	٤
١٢٠-		٢	٣	١	٦
١٥٠-			٣	٢	٥
المجموع	٢	٧	٧	٤	٢٠

المطلوب :

(أ) حساب الارتباط لقياس قوة العلاقة بين الظاهرتين .

(ب) أوجد معادلة انحدار الاستقطاعات على المرتب ثم قدر جملة الاستقطاعات عند المرتب ١٠٠ جنيه .

تمارين على الأرقام القياسية

١ – فيما يلي جدول يوضح أسعار وكميات مجموعة من السلع

السلعة	السنة ١٩٩٩		السنة ٢٠٠٤	
	سعر	كمية	سعر	كمية
أ	٥	١٥	٦	٢٠
ب	٦	٨	٧	١٢
ج	٧	١٢	٤	١٥
د	٤	٢٠	٥	١٨

والمطلوب إيجاد :

١- الرقم القياسي التجميعي البسيط للأسعار .

٢- الرقم القياسي للأسعار للاسبير .

٣- الرقم القياسي للأسعار لباش .

٤- الرقم القياسي للأسعار لفيشر .

٢ – فيما يلي بيان بالأسعار والكميات الخاصة بتكلفة أحد المنتجات الصناعية في إحدى الشركات خلال عامي ٢٠٠٠ و ٢٠٠٤ .

المنتجات الصناعية	الأسعار		الكمية	
	سعر	كمية	سعر	كمية
الألومنيوم	٧٠٠	١١٠٠	١٤٠	١٨٠
الحديد	١٠٠	١٥٠	١٢٠	١٥٠
الرصاص	٢٥٠٠	٢٠٠٠	٩٠٠	٦٠٠
الزنك	٩٠٠	٩٥٠	٤٠٠	٢٥٠
النحاس	١٠٠٠	١١٠٠	٦٠٠	٦٠٠

والمطلوب :

حساب الرقم القياسي التجميعي المرجح بالوسط الحسابي بكميات سنتي الأساس والمقارنة .

٣ - فيما يلي بيان بالأجور وعدد العمال في القطاعات المختلفة على المستوى القومي وذلك خلال عامي ٢٠٠٠ ، ٢٠٠٤ .

القطاع	متوسط الأجور		عدد العمال بالآلاف	
	٢٠٠٠	٢٠٠٤	٢٠٠٠	٢٠٠٤
الصناعات التحويلية	٤٠٠	٨٠٠	١٥٠	٢٠٠
الزراعة	٢٢٠	٣٥٠	١٨٠	١٥٠
التجارة	٧٠٠	٩٢٠	٧٠٠	٤٥٠
البناء والتشييد	٥٥٠	٧٤٠	٤٥٠	٧٤٠
النقل والمواصلات	٣١٠	٤٣٠	٣٨٠	٤٨٠

والمطلـــــوب :

تحديد مدى التغير الذي طرأ على أجور العمال في القطاعات المختلفـــة على المستوى القومي ما بين عامي ٢٠٠٠ ، ٢٠٠٤ وذلك باســتخدام الــرقم القياسي للوسط الهندسي لمناسيب الأجور مرجحاً بقيم سنة المقارنة .

٤ – فيما يلي جدول يوضح الأسعار والكميات لمجموعة من السلع :

السلعـــة	الســـــعر		الكميـــــــة	
	سنة ١٩٩٩	سنة ٢٠٠٤	سنة ١٩٩٩	سنة ٢٠٠٤
أ	٢٤٠	٣٦٠	٦	١١
ب	٤٠٠	٢٠٠	٥	١٣
ج	٣٠٠	٥٠٠	١٤	٢٨
د	١٨٠	٣٦٠	١١	٥

والمطلوب إيجاد :

١- الرقم القياسي التجميعي البسيط للأسعار .

٢- الرقم القياسي للأسعار المرجح بكميات سنة الأساس .

٣- الرقم القياسي للأسعار المرجع بكميات سنة المقارنة .

٤- الرقم القياسي الأمثل .

تمارين على الاحتمالات

١- صندوق به ٥ كرات سوداء ، ٤ بيضاء سحبت ٣ كـرات عشـوائياً ، أوجد احتمال أن يكون اثنان منهما سوداء .

٢- حقيبة بها ٣ كرات سوداء ، ٤ بيضاء ، وحقيبة أخرى بها ٥ كـرات سوداء وكرتان بيضاء ، نقلت كرة من الحقيبة الأولى إلى الثانيـة ثـم سحبت كرة من الحقيبة الثانية ، فما احتمال أن تكون الكرة المسـحوبة بيضاء ؟

٣- إذا كان ٢٥% من الطلبة ، ١٥% من الطالبات بإحدى الكليات يدرسون الرياضيات وكانت الطالبات تكون ٤٠% من العدد الكلي لطلبة الكلية. أختير أحد الطلاب عشوائياً ووجد أنه يدرس الرياضيات . فما احتمال أن يكون هذا الإختيار طالبة ؟

٤- ثلاث صناديق :

يحتوي الأول منها على ٣ كرات بيضاء ، ٤ كرات حمراء.

ويحتوي الثاني على ٣ كرات بيضاء ، ٥ كرات حمراء.

ويحتوي الثالث على ٢ كرة بيضاء ، ٣ كرات حمراء

تم إختيار صندوق عشوائياً وسحبت منه كرة عشوائياً .

(أ) فما احتمال أن تكون الكرة المسحوبة بيضاء ؟

(ب) إذا علم أن الكرة المسحوبة حمراء فما احتمـال أن تكـون مـن الصندوق الثالث ؟

المراجـــــــع

أولاً : مراجع باللغة العربية

١- فتحي محمد على ، الإحصاء التحليلي،مكتبة عين شمس ١٩٩٦/١٩٩٧.

٢- فتحي محمد علي ، مقدمة في علم الإحصاء، مكتبة عين شمس ١٩٧٤.

٣- هشام مخلوف ، عبد الله غالي ، عزت ندا ، أسس الإحصاء ، مكتبـة عين شمس ٢٠٠٤

٤- محمد جـلال الدين أبوالدهب ، هشام مخلوف ، محمد بدير السرجاني مبادئ الإحصاء ، مكتبة عين شمس ١٩٩٧/١٩٩٨ .

٥- أحمد حسن العطار ، مدخل إلى الإحصاء (محاضرات) ، معهد البحوث والدراسات الإحصائية – فرع الخرطوم – جامعة القاهرة .

٦- محمد صلاح الدين صدقي،مبادئ النظرية الإحصائية وتطبيقاتهـا فـي المشروعات التجارية والصناعية،الطبعة العاشرة،مكتبة عـين شـمس ١٩٩٩

٧- جـلال الصياد وآخرون ، مقدمة في الطـرق الإحصـائية وبحـوث العمليات ، الفاروق الحديثة للطباعة والنشر ١٩٨٦ ، ١٩٨٧ .

٨- محمد محمد عثمان ، مبادئ التحليل الإحصائي ، كلية التجارة ، جامعة الأزهر ، ٢٠٠٠ .

٩- محمد جبر المغربي ، عبد المنعم مرسى ، مبادئ الطرق الإحصائية ، جامعة المنصورة ، ١٩٩٠ .

١٠- حسن محمد حسين ، البحث الإحصائي ، دار النهضة ١٩٩٤ .

١١- فاروق عبد العظيم أحمد وآخرون ، الإحصاء ، دار الكتب الجامعية .

١٢- أحمد عبادة سرحان وآخرون ، مقدمة الإحصاء التطبيقي ، مكتبة دار المعارف ١٩٦٨ .

١٣- عبد الله عبدالحليم أبو بكر وآخرون ، الإحصاء ، دار التجارة والتعاون ، ١٩٨٣ .

١٤- سمير عبدا لمجيد ، علم الإحصاء الوصفي ، الرياض ، ١٩٩٥ .

١٥- أحمد ضياء الدين زيتون ، مبادئ علم الإحصاء ، كلية الزراعة جامعة المنيا ، ١٩٨٨

١٦- فتحي محمد علي وآخرون ، الإحصاء التجاري ، مكتبة عين شمس ١٩٩٥

١٧- محمد عبد السميع عناني ، الإحصاء (المبادئ والطرق) ، مكتبة حمادة ، الزقازيق ، ١٩٨٤ .

١٨- عثمان علي شلبي ، إبراهيم موسى عبد الفتاح ، أساليب التحليل الإحصائي ، مكتبة المدينة ، الزقازيق ١٩٨٥ .

١٩- سمير كامل عاشور ، سامية أبو الفتوح ، مقدمة في الإحصاء التحليلي ، دار الكتب ، القاهرة ، ١٩٨٧ .

٢٠- أحمد عمر ، مبادئ الطرق الإحصائية وتصميم التجارب ، كلية الزراعة ، جامعة الأزهر ، القاهرة ١٩٨٧ .

٢١- محمد أبو يوسف ، الإحصاء في البحوث العلمية ، المكتبة الأكاديمية ، القاهرة ، ١٩٨٩ .

٢٢- محمد توفيق المنصوري ، مصطفى عبد الغني أحمد ، الرياضية والتأمين ، مكتبة عين شمس ٢٠٠٣ .

ثانياً : المراجع الأجنبية

1- Steel, R.G.D. and J.H. Torrie (1980) : Principles and Procedures of Statistics. Megraw-Hill Kogakusha, LTD. Tokyo.

2- Ronald J. wonnacott & Thomas H. wonnacott, "Econometrics" , John Wiley and son Inc, 1970

3- Mode, E, "Elements of statistics", Prentice-Hall, 1961.

4- Clark, C.T. and Schkade, L.L. (1990): Statistical Methods for Business Decisions, South-Western Publishing Company Edward Arland LTD.

5- Yule, G, And Kendall, M.G, "Introduction to The Theory of Statistics" N.Y. Hafrer 1960.

6- Kuebler, R. R. & harrysmith, "Statistics, A Begining". John Wiley and Son, Inc., 1976.

7- Fisher, I., "the making of index Number Boston, H. Miffin 1963.

8- Crawshaw, J. and J. Chambers (1987) : A concise Course in A-Level Statistics Stanly Thorness (Publisher) LTD.

9- Mudget, B.D. Index Numbers" N.Y. willey 1971.

10- Gupta, C.B. (1973) : An Introduction to Statistical Methods Vikas Publishing House PVT LTD .

11- Skane, D.H. (1985) : Elementary Statistics, Addison-Wesley .

12- Freund, J.E. (1987) Modern Elementary Statistics, Prentice-Hall.